High-Throughput Hacking of the Methylation Patterns in Breast Cancer by In vitro Transcription and Thymidine-Specific Cleavage Mass Array on MALDI-TOF Silico-Chip

Ramin Radpour,1 Mahdi Montazer Haghighi,2 Alex Xiu-Cheng Fan,1 Peyman Mohammadi Torbati,3 Sinuhe Hahn,1 Wolfgang Holzgreve,1 and Xiao Yan Zhong1

1Laboratory for Prenatal Medicine and Gynecologic Oncology, Women’s Hospital/Department of Biomedicine, University of Basel, Basel, Switzerland; 2Department of Genetics, Azad University, East Tehran Branch; and 3Department of Pathology, Shaheed Beheshti Medical University, Tehran, Iran

Abstract

Over the last decade, the rapidly expanding interest in the involvement of DNA methylation in developmental mechanisms, human diseases, and malignancies has highlighted the need for an accurate, quantitative, and high-throughput assay. Existing methods are limited and are often too laborious for high-throughput analysis or inadequate for quantitative analysis of methylation. Recently, a MassCLEAVE assay has been developed using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry to analyze base-specific methylation patterns after bisulfite conversion. To find an efficient and more cost-effective high-throughput method for analyzing the methylation profile in breast cancer, we developed a method that allows for the simultaneous detection of multiple target CpG residues by using thymidine-specific cleavage mass array on matrix-assisted laser desorption/ionization time-of-flight silicon chips. We used this novel quantitative approach for the analysis of DNA methylation patterns of four tumor suppressor genes in 96 breast tissue samples from 48 patients with breast cancer. Each individual contributed a breast cancer specimen and corresponding adjacent normal tissue. We evaluated the accuracy of the approach and implemented critical improvements in experimental design. (Mol Cancer Res 2008;6(11):1702–9)

Introduction

DNA methylation is an important potential biomarker in cancer study and opens a new area in cancer therapy. Although DNA methylation analysis is a rapidly developing field, studies about its clinical usefulness are limited due to the fact that no single technique is superior (1–4). An accurate, sensitive, and reproducible high-throughput quantification of DNA methylation with compatibility of automation remains challenging (4).

A novel EpiTYPEr assay for high-throughput analysis of DNA methylation patterns using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has recently been introduced (5, 6). This assay is a tool for the detection and quantitative analysis of DNA methylation using MALDI-TOF MS and MassCLEAVE reagent, which enables base-specific (C/T) cleavage reactions (5, 7, 8). In a completely new and comprehensive study, by quantitative analysis of methylation patterns in a set of >400 candidate genes in 59 different cancer cell lines, the developer company showed robustness of (C/T) cleavage reactions in methylation studies (6).

The method uses T7-promoter–tagged PCR amplification of bisulfite-converted DNA, followed by generation of a single-stranded RNA molecules (9) and a subsequent procedure of base-specific cleavage (3’ to rUTP and rCTP) using RNase A (10-12). Bisulfite treatment of genomic DNA converts unmethylated cytosine into uracil while methylated cytosine remains unchanged. These C/T appear as G/A variations in the cleaved products generated from the reverse strand by C/T-specific cleavage. In the C-cleavage reaction, methylated regions are cleaved at every C to create fragments containing at least one CpG site each. For the T-cleavage reaction, both methylated and unmethylated regions are cleaved at every T to produce fragments (Fig. 1). These G/A variations result in a mass difference of 16 Da per CpG site, which is easily detected by the MassARRAY analyzer compact. In the mass spectrum, the relative amount of methylated sequence can be calculated by comparing the signal intensity between the mass signals of methylated and unmethylated regions.

We present an approach that allows for reduced costs based on the above system for high-throughput DNA methylation analysis and our modified protocol. We performed a T-specific cleavage reaction on CpG islands after bisulfite conversion of target sequences to analyze the methylated sequences. The relative amounts of methylated DNA were automatically
calculated by the software. We applied this technology for high-throughput quantification of methylation patterns of four tumor suppressor genes (Cadherin1, P16, RAR-b, and RASSF1; Supplementary data 1) in the paired cancerous and adjacent normal breast tissues from 48 patients with breast cancer. The genes were observed to be involved in cell adhesion, cell interaction, or gene expression during cancer development and progression.

FIGURE 1. A. MALDI-TOF MS DNA methylation analysis. Overview of the MassCLEAVE assay. Genomic DNA is bisulfite treated and PCR tagged to include the T7-promoter sequence. As shown, either top or bottom strand can be used for amplification. Subsequent shrimp alkaline phosphatase treatment, in vitro transcription using T7R&DNA polymerase and a specific nucleotide mixture plus RNase A digestion results in specific fragmentation. The obtained mixture of fragments can be analyzed by MALDI-TOF-MS. B. In our newly developed method, we eliminated the C-cleavage, and by using T-cleavage we achieved the same result. In this method, spectra 1 and 3, which were related to C-cleavage, were deleted; spectrum 2 was related to the T-cleavage of unmethylated DNA; and spectrum 4 was related to the T-cleavage of methylated DNA.
Results

Assay Development

The methylation quantification methodology presented here is schematically outlined in Fig. 1. In essence, the thymidine-specific reactions were realized by enzymatic cleavage of transcripts in which one specific base is present in the 2'-deoxy form (Table 1). In practice, not all spectral changes could be observed, and some fragments that were either too small or too large to be detected fall outside the utilizable mass window. Information may also be lost in the high-mass region. The length of the fragment may also influence the result of detection. We extended the utilizable mass range from 1,000 to <11,000 Da to increase the accuracy of detection. In the region <1,000 Da, we considered peaks detectable only when separated by ≥5 Da. Fragments that fall within this window and range in size from 4-mer to <30-mer can be calculated to cover ~76% of the target sequence. The use of a reference genomic sequence allows the resolution of remaining ambiguities spectra after reconstruction of sequence candidates from the cleavage pattern.

Accuracy of the Approach

We performed a number of experiments to confirm the accuracy of our modified T-cleavage assay for quantifying the methylation rate of the four tumor suppressor genes in breast cancer. In our design, a completely unmethylated template results in an RNA transcript that does not have any cleavable nucleotides and, hence, does not generate any cleavage products. A fully methylated template will result in an RNA transcript that contains several cleavable nucleotides and, thus, generate defined cleavage patterns. This distinct difference allows rapid and sensitive discovery of methylation sites.

We used the unique mass signals representing methylation events to test the specificity for detecting methylated DNA in mixed components. We mixed the fully methylated DNA of our designed positive control with the pure unmethylated DNA of our designed negative control using the sequence of tumor suppressor gene P16 in ratios of 100:0; 50:50; 25:75; 5:95, and 0:100. Figure 2A shows the resulting spectra of the T-specific cleavage mixture on MALDI-TOF MS. The assay was able to discriminate the methylated and unmethylated components according to the designed ratios.

Sensitive Detection

To evaluate the robustness and sensitivity of the T-cleavage assay in this study, we prepared a dilution series of a DNA on the promoter region of P16 gene, ranging from 80 to 2.5 ng DNA, to find the minimal input of target template necessary for a successful methylation analysis after in vitro transcription (Fig. 2B). The assay enabled the reliable detections of the methylated DNA in as little as 5 ng per PCR reactions. At the lowest template input of 2.5 ng, some of the replicates failed to generate detectable fragments with distinctive signal-to-noise ratios in the spectrum.

High-Throughput Quantitative Discovery of Methylated Sites in Breast Cancer

We applied this T-cleavage assay in the analysis of breast cancer patients. We analyzed the methylation patterns of four...
tumor suppressor genes (Cadherin1, P16, RAR-b, and RASSF1) in 96 cancerous and normal breast tissue samples from 48 patients with breast cancer. All of the CpG sites were analyzed by our T-cleavage assay using MALDI-TOF MS. The analyzed regions of the four tumor suppressor genes contained 71 CpG sites (total of 6,816 sites in 96 analyzed samples). The T-cleavage assay was able to detect >71% of CpG sites in amplicon (64% in CDH1, 74% in P16, 67% in RAR-b, and 78% in RASSF1; Table 2; Fig. 3). Nearly 44% of CpG sites were methylated at a very low degree, with average methylation <30%, and only 18% CpG units had mean methylation levels >90%. We performed two-way hierarchical clustering of the CpG unit methylation and the combined tumor and normal tissues in the training set to explore any natural groupings in this data set (Fig. 3). Using the cluster analysis, whereas high degrees of methylation in the CpG sites. For this reason, the C-cleavage will not be as informative for CpG-rich DNA, such as CpG islands. Therefore, only the T-cleavage assay was used in this study. This assay generates a defined experimental mass signal pattern wherein each mass signal represents at least one fragment of CpG site evolved from the target sequence. For analysis, this experimental pattern is subsequently compared with an in silico reference mass signal pattern derived from a reference genomic sequence (5). Differences between the expected and the observed mass signal pattern were interpreted, enabling quantitative analysis of methylated and unmethylated sequences. Finally, the data were compared with the original sequence. By using T-cleavage assay, the cost of this experiment can be reduced by half.

In our study, we assessed the accuracy of the T-cleavage assay by using positive (fully methylated DNA) and negative (unmethylated DNA) controls. The sensitivity and specificity of the assay were compared and confirmed by bisulfite sequencing analysis. The T-cleavage assay provides a rapid, automated discovery of multiple methylated CpG sites in regions of 200 to 600 bp and does not require cloning of PCR products for sequencing analysis.

Using the high-throughput robust T-cleavage assay on MALDI-TOF MS, we quantitatively analyzed methylation patterns of four tumor suppressor genes in 96 breast cancerous and adjacent normal tissues from patients with breast cancer. Hierarchical clustering identified substantial differences in the quantitative methylation profiling of tumor tissue compared with adjacent normal tissue (Fig. 1). We identified CpG units from the promoter regions of two genes (P16 and RASSF1) that exhibited significantly different levels of methylation (P < 0.001) between normal and tumor samples. Our analysis shows that RASSF1 is highly methylated in breast cancer tissue, and most significant methylation differences in this study were observed in genomic regions with relatively highly CpG density. Our data suggest that the high frequencies of methylation on the two tumor suppressor genes (P16 and RASSF1) in breast cancer tissues may serve as biomarkers for the disease.

In this study, the large-scale analysis is the first high-throughput implementation of this method for quantification

### Table 2. High-Throughput Methylation Analysis of Informative CpG Sites in 48 Breast Cancers

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amplicon Size (bp)</th>
<th>Total No. of CpG Sites in Amplicon</th>
<th>No. of Analyzed CpG Sites in Amplicon</th>
<th>No. of Analyzed CpG Sites in 48 Samples</th>
<th>% of CpG sites Missed Due to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total Sites</td>
<td>Single Sites</td>
<td>Combined Sites</td>
<td>Adjacent peak</td>
</tr>
<tr>
<td>Cadherin1 (CDH1)</td>
<td>422</td>
<td>14</td>
<td>9</td>
<td>192</td>
<td>0</td>
</tr>
<tr>
<td>P16 (CDKN2A)</td>
<td>459</td>
<td>27</td>
<td>20</td>
<td>768</td>
<td>1</td>
</tr>
<tr>
<td>RAR-b</td>
<td>445</td>
<td>15</td>
<td>10</td>
<td>288</td>
<td>3</td>
</tr>
<tr>
<td>RASSF1</td>
<td>382</td>
<td>41</td>
<td>32</td>
<td>1,020</td>
<td>3</td>
</tr>
</tbody>
</table>

**NOTE:** The in silico digestion was done for the T-cleavage assay. The percentage of informative CpG sites (total sites) in the amplicon is divided into single CpG sites (single site) and combined sites when adjacent CpG sites fall within one fragment or when fragment masses are overlapping. The percentage of CpG sites that give no methylation information due to interfering neighboring peaks or partially overlapping peaks (adjacent peak) or that will be missed because they fall outside of the usable spectral range (out of range) is shown.

### Discussion

In this study, we present a quantitative high-throughput assay with reduced costs for methylation analysis. On the MALDI-TOF MS, mass spectrum information can be used to determine methylated and unmethylated DNAs to assess the degree of methylation for each CpG island independently, and to estimate the average methylation for the entire target region (5, 6, 13, 14). Using this method, both hypermethylation and hypomethylation can be detected in samples.

Recently, the company developed a product that showed robustness of (C/T) cleavage with the MassARRAY system for methylation analysis based on a study assessing a set of >400 candidate genes in 59 different cancer cell lines (6). In the MassCLEAVE kit (Sequenom, Inc.), the T-cleavage reagent contains dCTP, rUTP, rGTP, and rATP, and the C-cleavage reagent contains a mixture of dTTP, rCTP, rGTP, and rATP, resulting in a unique cleavage 3’ of rTTP or rUTP only, respectively. In EpiTYPER two cleavage reactions, the reverse strand is cleaved by RNase A at specific bases [uracil (U) or cytosine (C)]. Cleavage products are generated for the reverse transcription reactions for both U [thymidine (T)] and C in separate reactions. In the C-cleavage reaction, methylated regions are cleaved at every C to create fragments containing at least one methylated CpG site. For the T-cleavage reaction, both methylated and unmethylated regions are cleaved at every T to produce fragments containing methylated and unmethylated CpG sites. Therefore, the T-cleavage assay was used in this study. This assay generates a defined experimental mass signal pattern wherein each mass signal represents at least one fragment of CpG site evolved from the target sequence. For analysis, this experimental pattern is subsequently compared with an in silico reference mass signal pattern derived from a reference genomic sequence (5). Differences between the expected and the observed mass signal pattern were interpreted, enabling quantitative analysis of methylated and unmethylated sequences. Finally, the data were compared with the original sequence. By using T-cleavage assay, the cost of this experiment can be reduced by half.

In our study, we assessed the accuracy of the T-cleavage assay by using positive (fully methylated DNA) and negative (unmethylated DNA) controls. The sensitivity and specificity of the assay were compared and confirmed by bisulfite sequencing analysis. The T-cleavage assay provides a rapid, automated discovery of multiple methylated CpG sites in regions of 200 to 600 bp and does not require cloning of PCR products for sequencing analysis.

Using the high-throughput robust T-cleavage assay on MALDI-TOF MS, we quantitatively analyzed methylation patterns of four tumor suppressor genes in 96 breast cancerous and adjacent normal tissues from patients with breast cancer. Hierarchical clustering identified substantial differences in the quantitative methylation profiling of tumor tissue compared with adjacent normal tissue (Fig. 1). We identified CpG units from the promoter regions of two genes (P16 and RASSF1) that exhibited significantly different levels of methylation (P < 0.001) between normal and tumor samples. Our analysis shows that RASSF1 is highly methylated in breast cancer tissue, and most significant methylation differences in this study were observed in genomic regions with relatively highly CpG density. Our data suggest that the high frequencies of methylation on the two tumor suppressor genes (P16 and RASSF1) in breast cancer tissues may serve as biomarkers for the disease.

In this study, the large-scale analysis is the first high-throughput implementation of this method for quantification
FIGURE 3. High-throughput analysis of informative CpG sites for the four studied tumor suppressor genes (A-D). The color of circles is related to percent of methylation in each CpG site. Arrows show the different methylation pattern between cancerous tissues and adjacent normal tissues (left pictures). Right pictures are two-way hierarchical cluster analysis of 48 cancerous breast tissues and 48 normal breast tissues.
of methylation in breast cancer. We were able to conclude that T-specific assay in combination with MALDI-TOF MS is a sensitive, accurate, and reliable technique for cost-effective high-throughput methylation analysis. The T-cleavage assay is readily amenable to automation. Our results support the suitability and robustness of T-cleavage assay on MALDI-TOF MS for cancer studies as well as for understanding methylation events in life science.

Materials and Methods

Samples

The study was approved by the local institutional review board. DNA samples from human hepatocyte cells were used to develop a methylation-positive control and human genomic DNA samples were used to develop a methylation-negative control for the assessment of the sensitivity and specificity of the approach.

Samples were collected from 48 paraffin-embedded cancerous breast tissues and paired adjacent normal tissues. Paraffin-embedded sections from adjacent normal and cancerous breast tissue samples were examined by two experienced pathologists. DNA was extracted from three to five sections of each 10-μm-thick paraffin-embedded sample (around 0.01-0.02 g of tissue) using a High Pure PCR Template Preparation Kit (Roche Diagnostics) and eluted into 150 μL of elution buffer. The eluted DNA was stored at −20°C until further use.

Bisulfite Treatment

To perform bisulfite conversion of target sequence, the EZ-96 DNA Methylation Kit (Zymo Research) was used according to the instruction manual. DNA was extracted from three to five sections of each 10-μm-thick paraffin-embedded sample (around 0.01-0.02 g of tissue) using a High Pure PCR Template Preparation Kit (Roche Diagnostics) and eluted into 150 μL of elution buffer. The eluted DNA was stored at −20°C until further use.

Primer Design and PCR Tagging for EpiTYPER Assay

CpG density and CpG sites of target sequences of the four tumor suppressor genes were analyzed with extreme precision for the PCR primer design. We designed primers using MethPrimer (15). For sequences that exceeded 450 bp in

FIGURE 4. A. Primers for in vitro transcription. A-1, reverse primer with T7-promoter tag. A-2, forward primer with 10-mer tag sequence as balance. B, PCR product cloning of P16 tumor suppressor gene into T-vector as positive or negative control. C, Bisulfite-treated sequence after treatment with methylationtransferase (M.SssI) enzyme as positive control. D, Bisulfite-treated sequence after first round of PCR as negative control. E, EcoRI restriction enzyme digested a 459-bp DNA amplified from P16 tumor suppressor gene after bisulfite conversion. Lane 1, DNA treated with Bisulfite Kit and cytosine converted to thymine; therefore, it made a recognition site for EcoRI at position 379. Lane 2, in nontreated DNA, there was no recognition site for enzyme.
length, we randomly assigned a 450-bp stretch within the annotated CpG islands. In PCR amplification, a T7-promoter tag was added to the reverse primer and a 10-mer tag sequence to the forward primer to balance the PCR primer length (Fig. 4A). The primer sequences, annealing temperatures \( T_a \), and PCR conditions are described in Supplementary data 2.

**Synthesized Fully Methylated DNA as Positive Control and Unmethylated DNA as Negative Control**

(a) To set up positive control, DNA sample from human hepatocyte cells was treated with the CpG methyltransferase (M.SssI) enzyme. M.SssI was isolated from a strain of *E. coli*, which contains the methyltransferase gene from *Spiroplasma* sp. strain MQ1 (New England Biolabs) and methylated all cytosine residues (C5) within the double-stranded dinucleotide recognition sequence 5’-CG-3’. S-Adenosylmethionine was used as a source of methyl groups. After treatment with M.SssI, DNA was treated with bisulfite conversion.

(b) To set up negative controls, we amplified genomic DNA using selected primer pairs before bisulfite conversion (because there was no methylase enzyme in the PCR reaction). Afterward, PCR products were converted by the bisulfite treatment, which changed whole cytosines into thymidines.

The sequence accuracy of the designed positive control was confirmed by direct bisulfite sequencing, and the sequence accuracy of the negative controls was checked by bisulfite sequencing after cloning PCR products into T-vector. Clones were sequenced from the pooled PCR reactions using the Dye Terminator cycle sequencing kit with AmpliTaq DNA polymerase FS (Applied Biosystems) and the automated 373A NA Terminator cycle sequencing kit (Sequenom) according to the instruction manual. After treatment with M.SssI, DNA was treated with bisulfite conversion.

**Dephosphorylation of Unincorporated Deoxynucleotide Triphosphates by Shrimp Alkaline Phosphatase**

Following the PCR amplification, 5 µL of PCR products were mixed with 2 µL of shrimp alkaline phosphatase enzyme (Sequenom) to dephosphorylate unincorporated deoxynucleotide triphosphates from the PCR reactions. The mixtures were incubated at 37°C for 20 min and then at 85°C for 5 min to inactivate the shrimp alkaline phosphatase enzyme.

**In vitro Transcription and RNase A Cleavage**

In the first step, transcription/RNase A cocktail for T-cleavage reaction was prepared using MassCLIVE (hMC) kit (Sequenom) according to the instruction manual. After shrimp alkaline phosphatase treatment, 2.5 µL of the mixture were directly added to a 5 µL transcription cocktail, containing 20 units of T7 R&DNA polymerase (Epicentre) to incorporate dTTP in the transcripts. Ribonucleotides were used at 1 mmol/L and the deoxynucleotide triphosphate substrate at 2.5 mmol/L; other components in the reaction were as recommended by the supplier. Following the *in vitro* transcription, RNase A was added to cleave the *in vitro* transcripts. In the cleavage reaction, the reverse strand was cleaved by RNase A at specific bases (U).

The mixture of cleaved fragments was further diluted with double-distilled water to a final volume of 27 µL. The phosphate backbone, before MALDI-TOF MS, was cleaned by adding 6 mg CLEAN Resin (Sequenom).

**Mass Spectrometry**

Twenty-two nanoliters of cleavage reaction were robotically dispensed (nanodispenser) onto silicon chips preloaded with matrix (SpectroCHIP, Sequenom). Mass spectra were collected using a MassARRAY mass spectrometer (Sequenom). Depending on the sequence of the target region and the distribution of CpGs, the mass spectrum may contain multiple signal pairs of cleavage products. These signal pairs can be used to estimate the ratio of methylated to unmethylated DNA. Finally, the results and spectra were analyzed with EpiTYPER v.1.0 software.

**Statistical Methods**

The two-way hierarchical cluster analysis clustered the 96 tissue samples (48 tumor derived and 48 paired normal tissues) and most variable CpG fragments for each gene based on pairwise Euclidean distances and the complete linkage clustering algorithm. The method involves first establishing the strength of the connection between two samples (called distance), followed by the reorganization of the samples according to their relationship to each other. The algorithm “clusters” samples divided into same clusters according to their similarity, and the double dendrogram was used to visualize the results. The method presented in this article clustered samples along the x axis and CpG units along the y axis. The procedure was carried out using the double-dendrogram function of the Gene Expression Statistical System for Microarrays (GESS) version 7.1.5 (NCSS).

**Additional Information**

The list and complete sequences of CpG islands in selected tumor suppressor genes, the bisulfite converted sequences, the sequences of PCR-tagged primers for *in vitro* transcription, and the PCR conditions are available in supplementary data.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**References**

High-Throughput Hacking of the Methylation Patterns in Breast Cancer by In vitro Transcription and Thymidine-Specific Cleavage Mass Array on MALDI-TOF Silico-Chip

Ramin Radpour, Mahdi Montazer Haghighi, Alex Xiu-Cheng Fan, et al.


Updated version
Access the most recent version of this article at:
http://mcr.aacrjournals.org/content/6/11/1702

Supplementary Material
Access the most recent supplemental material at:
http://mcr.aacrjournals.org/content/suppl/2008/11/14/6.11.1702.DC1

Cited articles
This article cites 15 articles, 9 of which you can access for free at:
http://mcr.aacrjournals.org/content/6/11/1702.full.html#ref-list-1

Citing articles
This article has been cited by 5 HighWire-hosted articles. Access the articles at:
/content/6/11/1702.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.