SEQUENOM®
Product Preview Note

Introduction to DNA Methylation Analysis Using the MassARRAY System

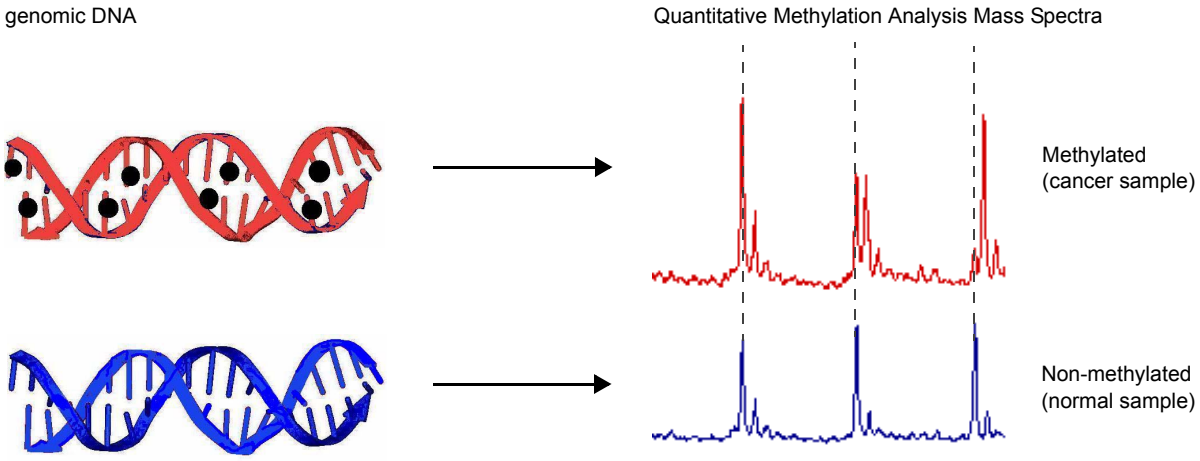
Mathias Ehrich, Devan Correll, and Dirk van den Boom

Introduction

Genetic information is not merely contained in the arrangement of the four nucleotide bases, but also in the covalent addition of methyl groups to cytosine within CpG dinucleotides. Methylation and related chromatin changes are important processes in the regulation of gene expression. The relevance of DNA methylation has been demonstrated in mammalian development, imprinting and X-chromosome inactivation[1], suppression of parasitic DNA[2], and various cancer types[3-6]. The ability to detect and to quantify methylation is particularly important to the field of cancer diagnostics. Changes in the methylation status of DNA have the potential to serve as an early detection marker for malignancies.

In the field of cancer research there is a need for an efficient method for performing quantitative DNA methylation analysis for a broad throughput range. Existing methods are limited; they are often too laborious for high throughput or inadequate for quantifying methylation. This article demonstrates that Quantitative Methylation Analysis is able to overcome the limitations apparent in existing methods. (See Figure 1.)

Quantitative Methylation Analysis is a tool for the discovery and quantitative analysis of DNA methylation that uses base-specific cleavage and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) introduced previously for SNP discovery[7-9]. Using the speed and accuracy of the MassARRAY® system, this method is ideal for discovery of methylation, for discrimination between methylated and non-methylated samples, and for quantifying the methylation levels of DNA.



The red-colored genomic DNA represents a methylated sample. Methylated CpG sites are represented by black circles. The blue-colored genomic DNA represents a normal sample. The mass spectra represent typical signal patterns seen for these types of samples after they have been processed using Quantitative Methylation Analysis. The dotted lines mark the signals representing non-methylated template DNA.

Figure 1. Example of Quantitative Methylation Analysis mass spectra indicating methylation levels in cancer and normal samples

Applying the Quantitative Methylation Analysis Assay

Overview

Quantitative Methylation Analysis is a bisulfite-treatment-based method for detection and quantitation of DNA methylation. Bisulfite treatment of genomic DNA converts non-methylated cytosine into uracil while methylated cytosine remains unchanged. Bisulfite treatment produces methylation-dependent sequence variations of C to T in the amplification products. These C/T variations appear as G/A variations in the cleavage products generated from the reverse strand by base-specific cleavage. These G/A variations result in a mass difference of 16 Da per CpG site, which is easily detected by the MassARRAY system. In the mass spectrum, the relative amount of methylation can be calculated by comparing the signal intensity between the mass signals of methylated and non-methylated template.

The method starts with bisulfite treatment of genomic DNA, followed by PCR amplification in which a T7-promoter tag is introduced. The PCR primers should be designed to yield a product within a 200-600 bp range. (Bisulfite treatment often limits the success of PCR when longer amplicons are used.) The significant advantage of this method is that the PCR primers are independent of the methylation state of the genomic DNA, meaning they bind to both methylated and non-methylated template (as opposed to methylation-specific primers). Only two primers are needed to screen for methylation changes within a region of several hundred bases in a single experiment. Next, *in vitro* RNA transcription is performed on the reverse strand, followed by base-specific cleavage. In the cleavage reaction, the reverse strand is cleaved by RNase A at specific bases (U or C). Cleavage products are generated for the reverse transcription reactions for both U (T) and C in separate reactions. (T is used from this point forward, since it is found in DNA. Cleavage, however, actually happens at U in an RNA molecule.) MALDI-TOF MS analyzes the cleavage products, and a distinct signal pair pattern results from the methylated and non-methylated template DNA. (See Figure 2.)

Quantitative Methylation Analysis generates quantitative results for each cleavage product analyzed. Each cleavage product encloses either one CpG site or an aggregate of multiple CpG sites. An analyzed unit containing one or multiple CpG sites is called a "CpG unit." For both T and C reactions, the resulting cleavage products have the same length and differ only in their nucleotide composition. Using this approach, Quantitative Methylation Analysis was applied to a model system using a CpG island in the IGF2/H19 region[10]. The resulting mass signal pattern represented hemi-methylated DNA, and the signal intensity ratio of methylated to non-methylated signals was equivalent to 0.5, as expected. The standard deviation in quantitation is 5%. (See Figure 3 and Figure 4 on page 3.)

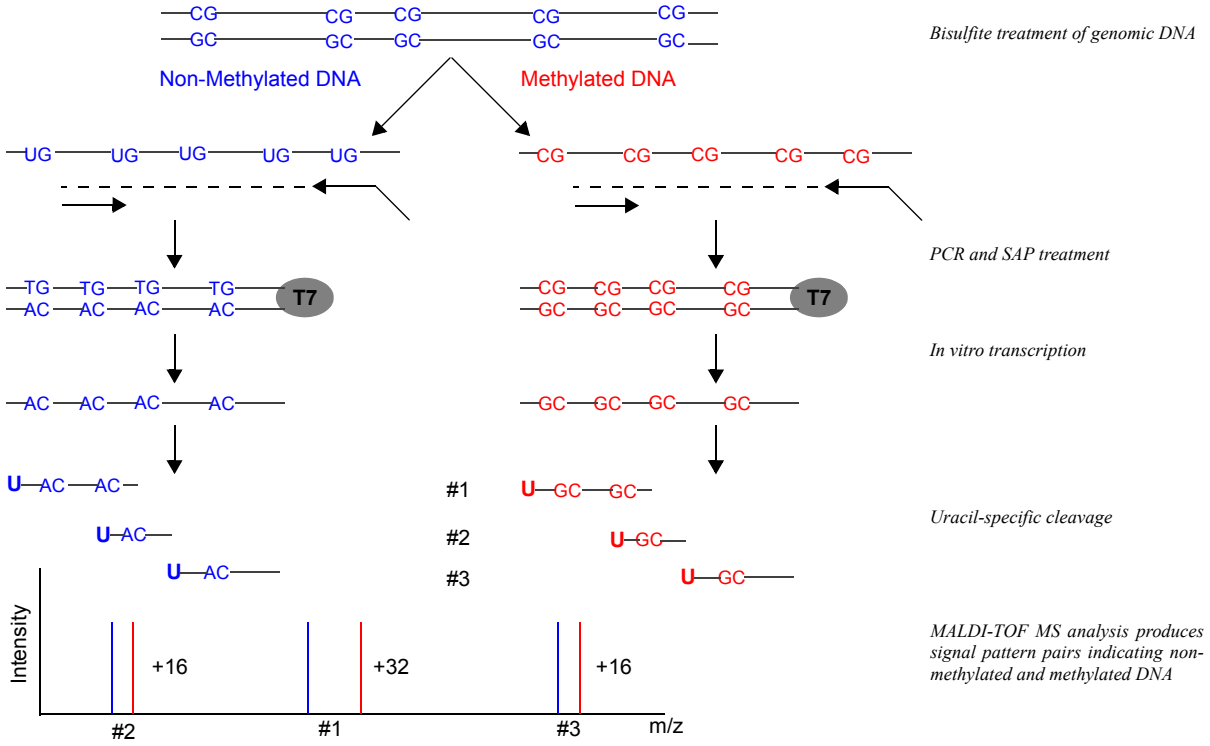
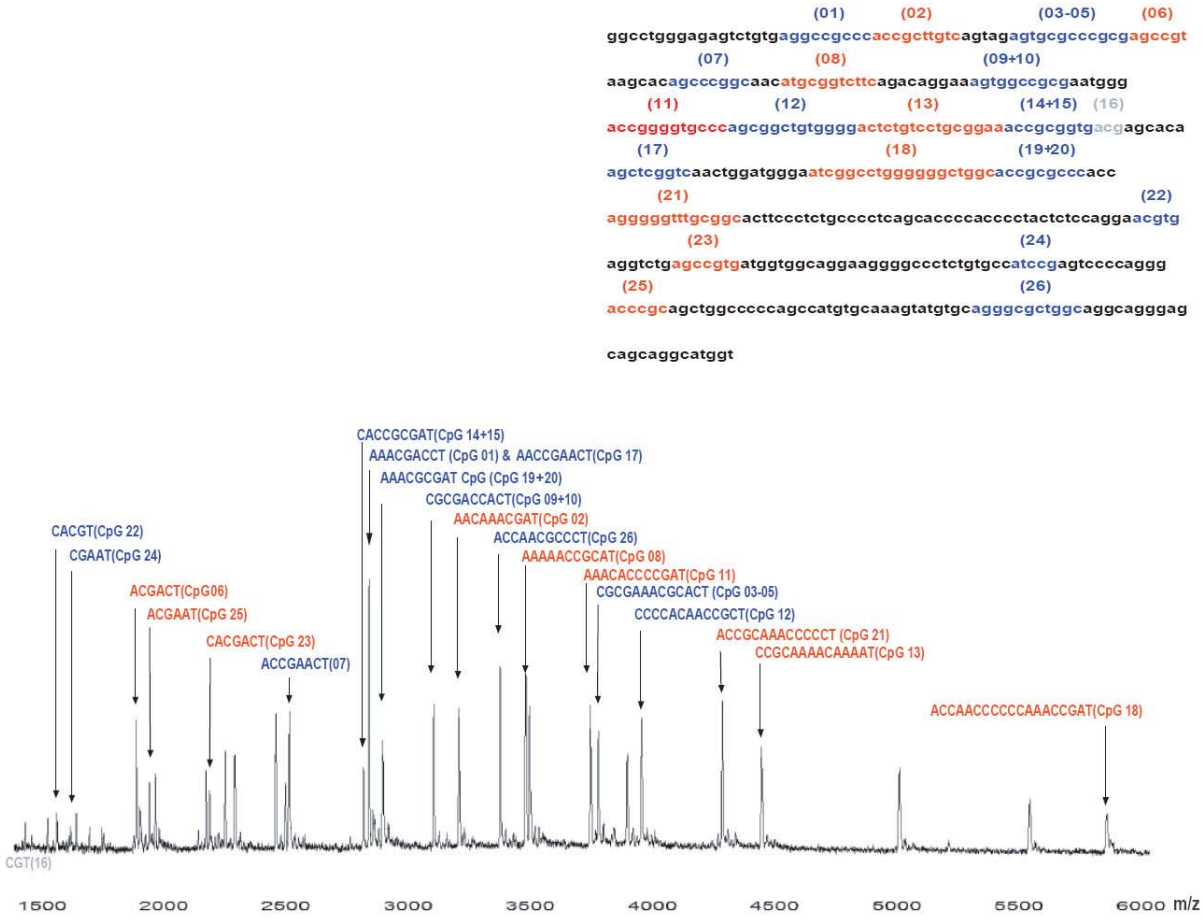


Figure 2. Overview of Quantitative Methylation Analysis process

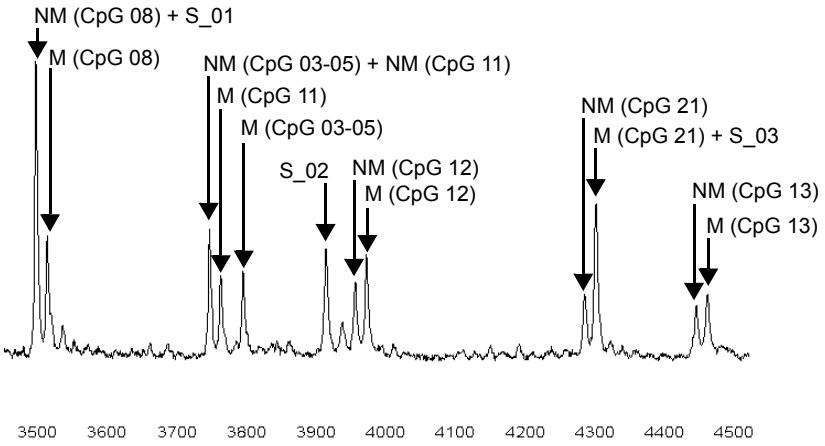


Example amplicon from genomic DNA (above right).

Example of mass spectra for T-reverse cleavage reaction and sequences of corresonding cleavage products from fully methylated DNA (above).

Each section of blue and red text used in the amplicon and spectra illustrations represents a cleavage product. The CpG numbers shown in parentheses correspond between the amplicon illustration and the spectra illustration.

Figure 3. Correlation between amplicon sequence and signal pattern for T-reverse cleavage reaction



Example of mass spectra generated by T-reverse cleavage reaction of the IGF2/H19 region. Arrows mark the position of mass signals representing non-methylated (NM) and methylated (M) CpGs. Signals with 16 Da shift (or a multiple thereof) represent methylation events.

Silent signals (S) are not affected by methylation but may overlay other signals. Their signal intensity changes can still be used to quantitate methylation.

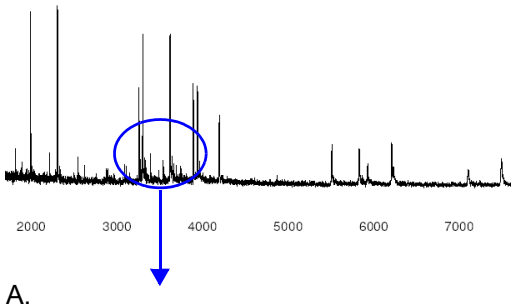
Figure 4. Spectra for hemi-methylated DNA, where relative methylation levels are 50%, as expected

Mass Spectrum Reveals Methylation Patterns

As mentioned earlier, bisulfite treatment introduces a C to T shift on the forward strand. This variation leads to a G to A shift on the reverse strand, which is represented in the mass spectrum by signal pairs separated by 16 Da (or multiples thereof when multiple fragments are enclosed). In the resulting mass spectrum, there are distinct signals illustrating non-methylated and methylated template DNA. (See Figure 5.) 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 determine the ratio of non-methylated to methylated DNA.

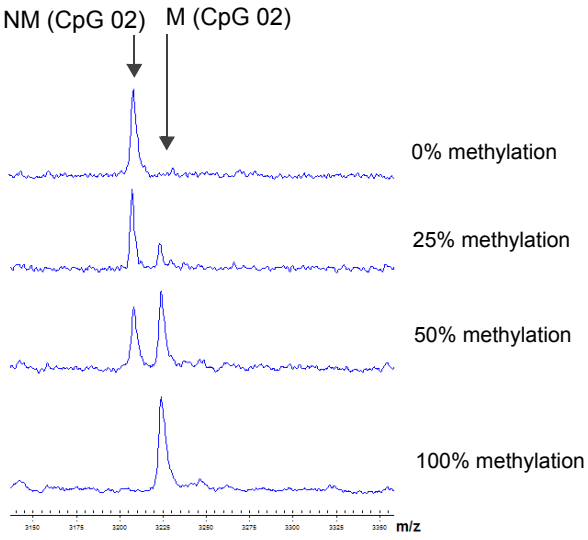
When a simulation was run for all 27,000 known CpG sites in the human genome, the Quantitative Methylation Analysis method was able to quantify approximately 90% of the CpG sites using two cleavage reactions. (Where CpG sites were longer than 500 bp, a 500 bp region was randomly selected for simulation.)

Information presented in the mass spectrum can be used to assess the degree of methylation for each CpG unit independently or to estimate the average methylation quantity for the entire target region. Using this method, both hypermethylation and hypomethylation can be detected in samples. The MassARRAY system is able to detect the methylation level in a mixture as low as 5%.



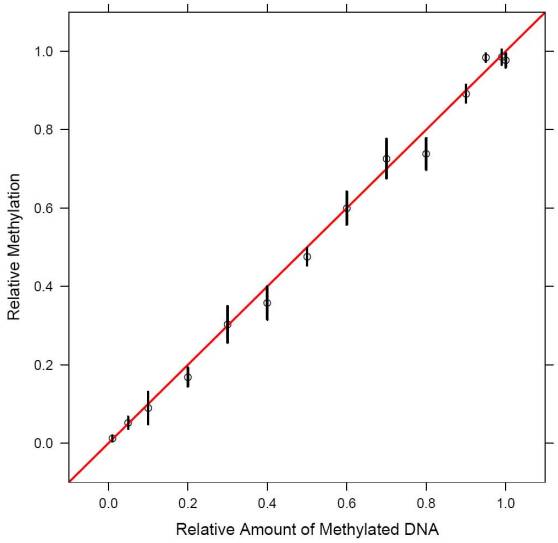
A.

Figure 5A illustrates signal patterns for a single CpG unit. It is an example illustration taken from a “zoomed in” view of a spectrum generated in Quantitative Methylation Analysis.



The non-methylated (NM) DNA fragment (5’-AACAAA**CA**AT-3’) is in the same position as the methylated. The methylated (M) DNA fragment (5’-AACAAA**CG**AT-3’) shows a mass signal that is 16 Da greater.

B.



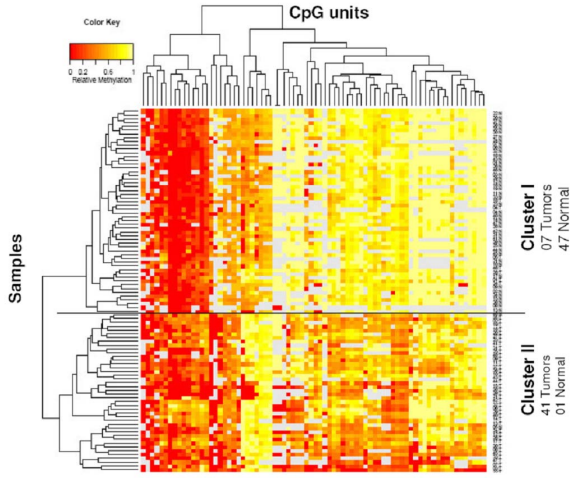
Graph showing the relationship between the estimated relative methylation and the relative amount of methylated DNA spiked into the mixture.

Figure 5. A.) Example of mass signal pairs (for one CpG unit) showing methylation ratios
B.) Graph of relative methylation levels

Large-Scale Application of Quantitative Methylation Analysis

In a study conducted with the Roy Castle International Centre for Lung Cancer Research, University of Liverpool, a large-scale application of Quantitative Methylation Analysis was used to characterize DNA methylation in 48 non-small cell lung cancer (NSCLC) patients. Forty-seven genes and their promoter regions were selected from literature and publicly available gene expression data. A total of 1,425 CpG sites were quantitatively analyzed for DNA methylation. Two-way hierarchical clustering was performed based on the quantitative methylation results. This clustering was used to identify differentially methylated regions, which contribute to the disease phenotype. The NSCLC study revealed significantly different methylation patterns between normal and tumor tissue.

Quantitative methylation data was collected for 599 CpG units. The relative methylation levels were calculated by comparing the signal intensity for the mass signals from methylated and non-methylated template DNA. The relative methylation data was filtered for quality (data must be available for more than 90% of all samples) and for variance (variance >0.01). The resulting subset of 248 CpG units was used in a two-way hierarchical clustering analysis. This analysis identified two main clusters that are mostly populated by either tumor or normal samples, indicating that Quantitative Methylation Analysis is useful in the precise characterization of epigenetic changes underlying disease phenotypes. See Figure 6.



This two-dimensional cluster was generated with the Heatmap.2 function from the g plots package using the statistics software R.

Figure 6. Cluster of relative methylation from NSCLC and normal samples

Highlights

Quantitative Methylation Analysis provides the following benefits:

- Rapid, automated discovery of multiple methylated CpG positions in regions of 200-600 bp
- Simplified assay design; no extension primers needed
- No need for cloning of PCR products
- Quantitative assessment of the degree of methylation
- Simple identification of hypomethylation and hyper-methylation
- Detection of methylation levels as low as 5% in sample mixtures
- Results may be obtained from various sample types, including frozen tissue, mouth swabs, and paraffin-embedded tissue
- High precision and reproducibility
- Scalable throughput on an established MassARRAY system
- Gene expression, SNP discovery, genotyping, and allelotyping studies may be run on same platform

Methods

Primer Design

Primers are designed as illustrated in Figure 7. The recommended size range for PCR amplicons is 200-600 bp. (Longer amplicons were also analyzed, but bisulfite treatment limits the success of PCR.) Primers were designed using MethPrimer (http://www.urogene.org/methprimer/).

Design the following primer systems for use in methylation analysis:

- T7-promoter tagged reverse primer to obtain an appropriate product for *in vitro* transcription. An 8 bp insert is included to prevent abortive cycling.
- 10mer-tagged forward primer to balance the PCR.

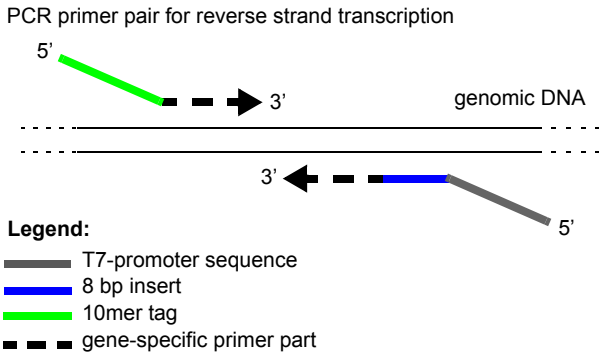


Figure 7. PCR primers for reverse strand transcription

Bisulfite Treatment

To run bisulfite conversion treatment, use either the EZ-96 DNA Methylation Kit or the EZ DNA Methylation Kit from Zymo Research (http://www.zymoresearch.com)

Step 1: Amplification

Amplify 1 µL DNA in a 5 µL volume using a 384-microtiter format. (Use 1.00 µL of at least 5 ng/µL DNA for a final concentration of 1 ng/µL per reaction.) Each PCR will be split into two cleavage reactions (T Cleavage reaction and C Cleavage reaction).

Seal the plates and cycle as follows:

- 94º C for 15 minutes
- 45 cycles:
  94º C for 20 seconds,
  62º C for 30 seconds
  72º C for 1 minute
- 72º C for 3 minutes

Note: If you have questions about the suitability of your genomic DNA or performing the bisulfite treatment, contact Sequenom customer support at 1 877 4 GENOME (toll free within the U.S.) or at helpdesk@sequenom.com.

Step 2: Dephosphorylation

Add 2 µL of Shrimp Alkaline Phosphatase (SAP) enzyme to each 5 µL PCR to dephosphorylate unincorporated dNTPs from the PCR. Incubate the plates for 20 minutes at 37º C. Then, incubate at 85º C for 5 minutes.

Step 3: In Vitro Transcription and RNase A Cleavage

Prepare transcription/RNase A cocktail for each cleavage reaction (T and C). For standard setup, prepare one transcription/RNase A cocktail per plate. Add 5 µL of transcription/RNase A cocktail and 2 µL of PCR/SAP sample into a new, uncycled microtiter plate. Centrifuge the plates for one minute. Then, incubate the plates at 37º C for 2 hours.

Step 4: Sample Conditioning

Add 20 µL of ddH_2O to each sample within the 384-well plate. Add 6 mg of Clean Resin to each well using the resin plate. Rotate for 10 minutes and spin down for 5 minutes at 3,200 x g.

Note: To avoid pH shifts, always add ddH_2O first before adding the Clean Resin.

Step 5: Sample Transfer

Dispense 10-15 µL of Quantitative Methylation Analysis reaction product onto a 384-element SpectroCHIP® bioarray.

Step 6: Sample Analysis

Acquire spectra from the two cleavage reactions using the MassARRAY system.

Step 7: Analysis Software

Results can be analyzed using the Quantitative Methylation Analysis software. This software is not yet available for purchase. Commercial release of Quantitative Methylation Analysis software is scheduled for fall 2005.

Quantitative Methylation Analysis Requirements

Hardware and Software

- MassARRAY Analyzer (part number 00450)
  or
  MassARRAY Analyzer Compact
  (part number 10900)

- Quantitative Methylation Analysis software
  (part number to be determined)

Consumables

- hMC Transcription Kit (part number 10064)

Contact your SEQUENOM sales or customer support representative for information.

References

1. Li, E., "Chromatin modification and epigenetic reprogramming in mammalian development." *Nat Rev Genet*, 2002. 3(9): p. 662-73.

2. Walsh, C.P., J.R. Chaillet, and T.H. Bestor, "Transcription of IAP endogenous retroviruses is constrained by cytosine methylation." *Nat Genet*, 1998. 20(2): p. 116-7.

3. Costello, J.F., *et al*., "Aberrant CpG-island methylation has non-random and tumour-type-specific patterns." *Nat Genet*., 2000. 24(2): p. 132-8.

4. Jones, P.A. and S.B. Baylin, "The fundamental role of epigenetic events in cancer." *Nat Rev Genet*, 2002. 3(6): p. 415-28.

5. Costello, J.F. and C. Plass, "Methylation matters." *J Med Genet*, 2001. 38(5): p. 285-303.

6. Feinberg, A.P., "Cancer epigenetics takes center stage." *Proc Natl Acad Sci U S A*, 2001. 98(2): p. 392-4.

7. Stanssens, P., et al, "High-Throughput MALDI-TOF Discovery of Genomic Sequence Polymorphisms." *Genome Research*, 2004. 14(1): p.126-133.

8. "SNP Discovery Using the MassARRAY System." Application note. Sequenom web site: http://www.sequenom.com/Assets/pdfs/appnotes/SNP_Discovery_Application_Note.pdf.

9. Ehrich, M., *et al*., "Multiplexed discovery of sequence polymorphisms using base-specific cleavage and MALDI-TOF MS." *Nucleic Acids Research*, 2005. 33(4): e38.

10. Vu, TH., *et al*., "Symmetric and assymetric DNA methylation in the human IGF2-H19 imprinted region." *Genomics*, 2000. 64(2): p. 132-143.

Trademarks

SEQUENOM, MassARRAY, MassEXTEND, and SpectroCHIP are registered trademarks of SEQUENOM, Inc.
MassCLEAVE is a trademark of SEQUENOM, Inc.

The polymerase chain reaction (PCR) described herein is a patented process. You must have a license to perform PCR amplification. **Purchase of MassARRAY system components and/or MassEXTEND reagents does not confer a license to perform PCR.** Information on purchasing licenses to practice the PCR process may be obtained by contacting the Director of Licensing at Applied Biosystems, Incorporated, 850 Lincoln Centre Drive, Foster City, California 94404 USA or at Roche Molecular Systems, Inc., 1145 Atlantic Avenue, Alameda, California 94501 USA.

All other trademarks or service marks set forth herein are the property of their respective owners.

Web Site www.sequenom.com

U.S. Toll Free 1 877 4 GENOME

Corporate Headquarters/Western US +1 858 202 9000

Eastern US +1 617 244 8777

Europe +49 40 899676 0

Asia Pacific +61 7 3845 3683

SEQUENOM®

April 15, 2005

Doc. No. 8876-007, R00
CO 050127