Journal of Cellular Biochemistry 88:138–143 (2003)

Oligonucleotide-Based Microarray for DNA Methylation Analysis: Principles and Applications

Huidong Shi,[1] Sabine Maier,[2] Inko Nimmrich,[2] Pearlly S. Yan,[1] Charles W. Caldwell,[1] Alexander Olek,[2] and Tim Hui-Ming Huang[1]*

[1]Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri School of Medicine, Columbia, Missouri 65203
[2]Epigenomics, AG, Berlin, Germany

Abstract Gene silencing via promoter CpG island hypermethylation offers tumor cells growth advantages. This epigenetic event is pharmacologically reversible, and uncovering a unique set of methylation-silenced genes in tumor cells can bring a new avenue to cancer treatment. However, high-throughput tools capable of surveying the methylation status of multiple gene promoters are needed for this discovery process. Herein we describe an oligonucleotide-based microarray technique that is both versatile and sensitive in revealing hypermethylation in defined regions of the genome. DNA samples are bisulfite-treated and PCR-amplified to distinguish CpG dinucleotides that are methylated from those that are not. Fluorescently labeled PCR products are hybridized to arrayed oligonucleotides that can discriminate between methylated and unmethylated alleles in regions of interest. Using this technique, two clinical subtypes of non-Hodgkin's lymphomas, mantle cell lymphoma, and grades I/II follicular lymphoma, were further separated based on the differential methylation profiles of several gene promoters. Work is underway in our laboratory to extend the interrogation power of this microarray system in multiple candidate genes. This novel tool, therefore, holds promise to monitor the outcome of various epigenetic therapies on cancer patients. J. Cell. Biochem. 88: 138–143, 2003.

Key words: DNA methylation; oligonucleotide microarray; CpG islands; cancer

Genetic alterations in the promoter and exonic sequences of a gene can result in silencing of the corresponding transcripts. Recent studies, however, have indicated an additional level of control of gene expression, termed epigenetics that does not involve changes of genomic sequences [Jones and Laird, 1999; Bird, 2002]. One prominent epigenetic alteration is DNA methylation, an enzymatic reaction bringing a methyl group to the 5th carbon position of cytosine located 5′ to guanosine in a CpG dinucleotide [Bird, 1986]. Across the human genome, ~80% of CpG dinucleotides are heavily methylated, but some areas remain unmethylated in GC-rich CpG islands that are 0.5–2 kb in length [Bird, 1986]. Almost half of the known genes have CpG islands in promoters and the first exon regions [Jones and Laird, 1999; Bird, 2002]. Methylation of CpG island loci is linked to transcriptional silencing of genes on the inactive X chromosome or genes nonessential for differentiated cells [Antequera et al., 1990; Lyon, 1999; Smiraglia et al., 2001].

In cancer cells, aberrant DNA methylation is frequently observed in normally unmethylated CpG islands and silences the function of normally expressed genes [Jones and Baylin, 2002]. If the silencing occurs in genes critical to growth inhibition, the epigenetic alteration could promote tumor progression [Jones and Baylin, 2002]. This silencing is achieved through the recruitment of repressor complexes and the remodeling of the local chromatin structure in the 5′ regulatory regions of genes, preventing the interaction of promoter with its

Grant sponsor: Epigenomics, Inc.; Grant sponsor: National Cancer Institute; Grant numbers: R21 CA-94441, RO1 CA-69065.

*Correspondence to: Tim Hui-Ming Huang, PhD, Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri, 115 Business Loop I-70 West, Columbia, MO 65203.
E-mail: huangh@health.missouri.edu

Received 5 August 2002; Accepted 6 August 2002
DOI 10.1002/jcb.10313

transcriptional machinery [Jones and Baylin, 2002]. More common than genetic alterations, the epigenetic-mediated silencing is seen in many tumor-suppressor genes and genes important for cell cycle function and DNA repair in many cancer types [Esteller et al., 2001].

The concept for a comprehensive analysis of CpG island hypermethylation becomes increasingly relevant because the methylation-mediated gene silencing in cancer can be pharmacologically reversed, leading to gene reactivation and potential tumor eradication [Widschwendter and Jones, 2002]. Toward this end, considerable progress in DNA microarray technologies has lead to new possibilities in epigenetic research. This article describes a recent innovative advance that uses oligonucleotide-based microarrays to analyze DNA methylation in a large number of candidate genes simultaneously [Adorján et al., 2002; Gitan et al., 2002]. We further demonstrate selected applications of this method along with future prospects of high-throughput DNA methylation analysis in cancer.

METHYLATION-SPECIFIC OLIGONUCLEOTIDE MICROARRAY

Figure 1 outlines the strategy of oligonucleotide-based microarray for DNA methylation analysis. In order to differentiate the methylated versus the unmethylated state of a CpG dinucleotide, genomic DNA from test samples is first treated with sodium bisulfite, which deaminates unmethylated cytosine to uracil, while methylated cytosine is resistant to this modification [Frommer et al., 1992]. Specific genomic regions of interest are then amplified by PCR, converting the modified UG to TG and conserving the originally methylated dinucleotide as CG [Frommer et al., 1992; Clark et al., 1994]. Primers are designed such that the sequences contain no CpG dinucleotides and are complementary to the flanking sequences of a 200–300-bp DNA target. This allows for an unbiased amplification of both methylated and unmethylated alleles by PCR. To streamline target preparation, bisulfite-treated DNA samples can be amplified in a 96-well format and the PCR products verified using a 96-well gel electrophoresis system. Also, multiplex PCR can be implemented to increase throughput. Target DNAs are then purified and labeled with Cy5 or Cy3 fluorescence dye for microarray hybridization.

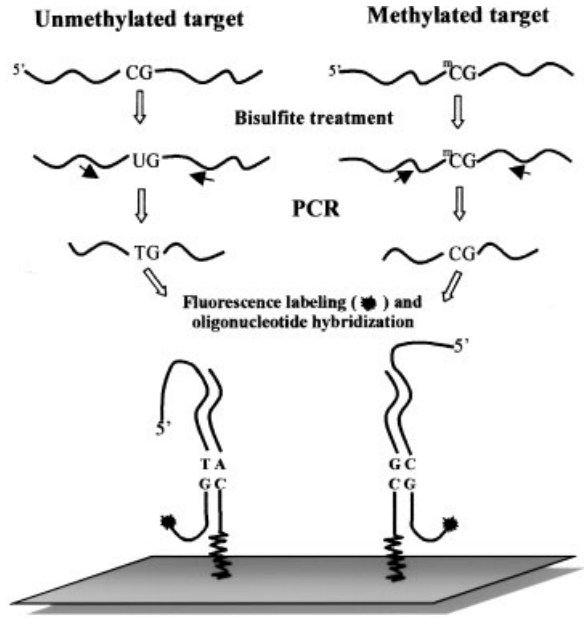

Fig. 1. Schematic outline for methylation-specific oligonucleotide microarray. Genomic DNA is bisulfite-treated and amplified by PCR for a specific CpG island region of interest. The amplified product is labeled with fluorescence dye and hybridized to oligonucleotide probes attached to a glass surface. At left an oligonucleotide probe is designed to form a perfect match with a DNA target containing the unmethylated allele. At right a probe is designed to form a perfect match with the methylated target.

Oligonucleotide probes affixed on solid support (e.g., microscope slides) via their 5′-ends are designed to form a duplex with methylated or unmethylated DNA targets. To insure that a probe will form a perfect match with a target, we used the Oligo 6^{TM} software (Molecular Biology Insights, Inc., Cascade, CO) and the following criteria to design oligonucleotide probes:

- Probes can be of various lengths between 17 and 23 nt.
- Each set contains a pair of probes, one for methylated (M) and the other for unmethylated (U) alleles.
- The interrogating CpG site(s) is preferably located in the center of a probe.
- As methylated CpG sites are present as palindromes, in the event where a probe for the forward strand is suboptimal for the assay, an oligonucleotide from the reverse direction can be used.
- DNA sequences with four consecutive T's, G's, A's, or C's are not used in the probe design.
- Selection of oligonucleotides with high melting temperature (50–60°C).

- Exclusion of oligonucleotides with the ability to form hairpin loops or secondary structures.

Each probe set is used to determine the methylation status of at least one CpG site. Sometimes, it is difficult to interrogate a single site due to high CpG density in a given sequence. In this case, probes are designed to interrogate two or more nearby CpG sites. This may limit our ability to detect methylation changes of single sites, but it is usually not necessary to define methylation of every CpG site within a CpG island in order to determine its effect on gene silencing. It is now known that the level of gene silencing is related to the overall density of CpG methylation in a promoter region [Stirzaker et al., 1997]. Alternatively, methylation of discrete regions in some promoter CpG islands is sufficient to invoke epigenetic silencing [Rice et al., 1998; Melki et al., 1999]. Hybridization conditions are selected to allow discrimination of nucleotide differences between M and U targets [Gitan et al., 2002]. The normal signal intensity of each hybridized probe is calculated and the ratio (M/M + U) for each probe set is derived. An example of this microarray assay is given in Figure 2.

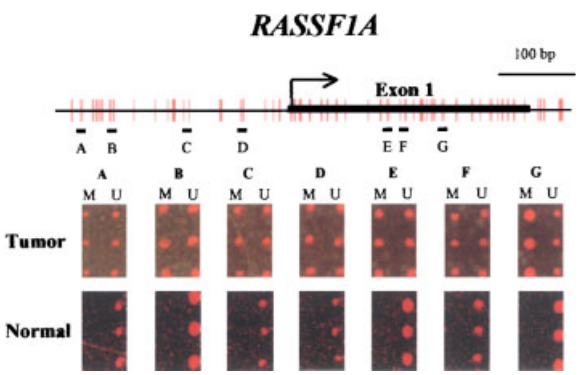

Fig. 2. Hybridization of breast tumor and normal samples to oligonucleotide microarray. The CpG dinucleotides (vertical bars) in the promoter and the first exon of *RASSF1A* (underscored by horizontal bars) are interrogated by unique oligonucleotides (M, methylated; U, unmethylated) to determine the methylation status of those sites. The transcription start site of *RASSF1A* is indicated by an arrow. Tumor and normal DNA samples are amplified after bisulfite treatment (see description in the text). Fluorescent dye, Cy5, is added to the 3′-end of amplified fragments via terminal transferase reaction. The labeled samples are hybridized to the microarray for 4 h at 50°C and washed prior to scanning. Hybridization images corresponding to the specified CpG dinucleotides are presented. Varying degrees of hybridization to both the ''M'' and the ''U'' oligonucleotides are observed in the tumor sample whereas only the ''U'' oligonucleotides light up in the normal sample. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Cross-hybridization likely occurs between imperfect-match probes and targets. Also, variations of the amounts between paired oligonucleotide probes printed on glass slides may occur. A unique control system is, therefore, implemented to test the sensitivity and specificity of the probes designed for microarray hybridization. The positive control is prepared by treating a DNA sample in vitro with *Sss*I methylase, an enzyme known to methylate all CpG's in the genome. The treated DNA is then subjected to bisulfite treatment and amplification by PCR. The target generated in this way is designated as 100% M (i.e., unconverted C's) for the specific CpG sites interrogated. Conversely, a PCR product of the same untreated genomic region is used as the negative control. The control target is expected to have 100% U (i.e., conversion of CG to TG) in an interrogating locus. The positive and negative controls are pooled in different proportions, representing methylation levels of 0, 25, 50, 75, and 100%. A series of experiments is conducted to define the range of CG:TG ratios that corresponds to each percentage of CpG methylation. It is expected that levels of DNA methylation for a given locus are proportional to the calculated intensity ratios [Gitan et al., 2002]. A standardization curve can be established and used to derive the level of methylation at the interrogating CpG sites. This approach can reliably detect differential methylation among different tumor samples [Gitan et al., 2002]. In theory, each probe should give 100% hybridization efficiency to differentiate M from U allele. However, this may not be the case in real experiments. In this case, the discriminative ability of each probe can be examined by the above control system and oligonucleotide pairs that have poor discriminative ability are discarded and redesigned.

METHYLATION PROFILING OF NON-HODGKIN'S LYMPHOMAS

As one example, using an oligonucleotide microarray platform previously described [Adorján et al., 2002], we determined DNA methylation profiles of the promoter CpG islands of 38 genes, mostly involved with cell-cycle regulation in non-Hodgkin's lymphomas [Caldwell et al., unpublished data]. Histological classification indicates that the disease entity exhibits remarkably heterogeneous subtypes.

Although not generally regarded as high-grade aggressive tumors, non-Hodgkin's lymphomas such as mantle cell lymphoma (MCL) and grades I/II follicular lymphoma (FLI/FLII), exhibit a spectrum of clinical behavior. We compared methylation patterns between MCL and FLI/FLII using oligonucleotide microarray. The measured values for the proportion of methylation (M/M + U) were calculated from six replicates, averaged across 40 test cases. A ranked matrix representation was then calculated using Wilcoxon rank statistics to identify differentially methylated loci that discriminate between 14 MCL and 26 FLI/FLII cases. Of the 38 genes initially analyzed, 16 genes ranked in decreasing order of significance ($P < 0.05$, corrected for multiple testing) are shown in Figure 3, with the least significant (*MYC*) at the top and the most significant (*TP73*) at the bottom. All of these were preferentially methylated in FLI/FLII, with the exception of *CDKN1C*. This methylation profiling served not only to gain a broader insight into differential gene methylation between these classes of NHL, but also set the stage for further experimental validation of the differential methylation of specific genes and their association with gene silencing. We found that the hypermethylated *AR* CpG island correlated inversely with its gene expression in these lymphomas (Caldwell, unpublished communication). Thus, differential methylation patterns in MCL and FLI/FLII discovered by this microarray assay provided further insights into methylation-associated gene silencing in non-Hodgkin's lymphomas.

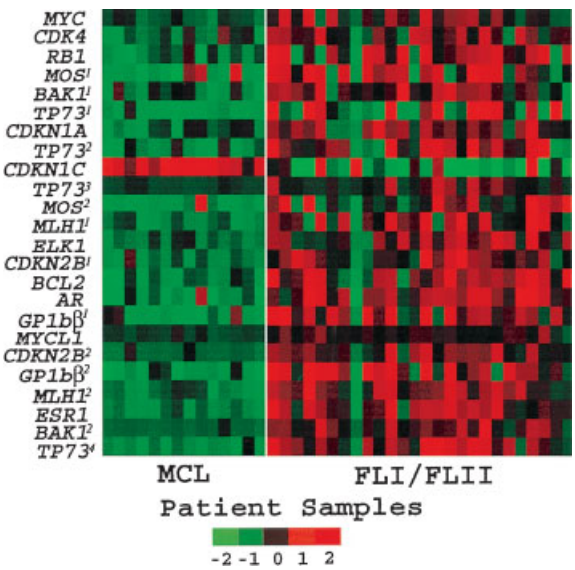

Fig. 3. A ranked matrix of differentially methylated CpG positions of individual gene loci between mantle cell lymphoma (MCL) and grades I/II follicular lymphoma (FLI/FLII). The microarray assay was conducted as described in the text. The Wilcoxon rank statistics were used to calculate corresponding *P* values for each CpG position. Gene names (loci number of each gene marked in superscript) labeled along the left Y-axis were all statistically significant at $P < 0.05$. The proportion of methylation was calculated for each locus, and across all cases based on values from six replicates (see description in the text). From this, the individual variation from the mean was calculated for each case. The methylation proportion relative to the mean of the group is expressed as red (increased methylation above the mean), green (decreased methylation below the mean), and black (no difference). The intensities of green versus red indicate the relative magnitude of the variation.

METHYLATION-BASED TUMOR CLASSIFICATION AND EPIGENETIC THERAPIES

Molecular markers have been used to classify tumors and group cancer patients according to clinical outcomes. However, this approach mainly uses a marker-by-marker approach, which is limited in throughput, to decipher complex diseases like cancer wherein a holistic view of the tumor genome becomes necessary. The recent development of high-throughput microarray technologies provides a powerful tool that can generate a vast amount of information in a single experiment. Gene expression profiles from cDNA microarrays have been used to differentiate different cancer types [Golub et al., 1999; Alizadeh et al., 2000; Clarke et al., 2001]. However, there are some practical problems with the large-scale microarray analysis of clinical specimens. As mRNA is the starting material, its unstable nature and the amount needed per array are some of the limitation for small biopsies. In this regard, profiles of altered DNA methylation provide an additional avenue for tumor classification. Mathematical tools such as hierarchical cluster analysis, support vector machines, and principle component analysis, used for expression profiling can also be applied for DNA methylation analysis [Model et al., 2001]. Adorján et al. [2002] used an oligonucleotide microarray to study the methylation profiles of lymphoblastic versus myeloid leukemias, and using supervised and unsupervised computation algorithms, they were able to predict and discover novel leukemia classes. Using the differential methylation hybridization technique, our laboratory found that profiles of altered DNA methylation effectively

classified breast cancer patients on the basis of their hormone receptor status and clinical aggressiveness [Yan et al., 2000, 2001], as well as predicted ovarian cancer patients with different progression-free survival [Wei et al., 2002]. The results of these studies enable future development of a new generation of methylation-based biomarkers.

The collective understanding of the role(s) of epigenetic control in cancer is just now starting to provide opportunities for translational research in this area. While DNA methylation is clearly an important mechanism, it is not the only epigenetic factor that modulates chromatin structures and gene function. Chromatin remodeling is a process that involves multiple mechanisms including DNA methylation, histone deacetylation, and methylation, and differential impact on the recruitment of transcriptional repressors [Jones and Baylin, 2002]. Clinical trials are now underway to treat various forms of cancer with demethylating agents, with or without added histone deacetylase inhibitors [Santini et al., 2001; Widschwendter and Jones, 2002]. In order to rationally predict responses to therapy, it is necessary to understand the broad range of methylation-associated changes in diseases, and perhaps equally important, to monitor the effects of these new pharmacological agents in patients under treatment. The microarray technology described herein may provide valuable information regarding the outcome of these epigenetic therapies. It is also likely that analysis of DNA methylation will become an important part of the cancer screening effort as we identify unique genes preferentially methylated in particular tumors. With this information, oligonucleotide microarrays can be tailor-made for the diagnosis and prognosis of different cancer types.

CONCLUDING REMARKS

Studies of the relationship of DNA methylation and gene silencing is a near ideal system because it is now known that multiple critical genes can become silenced through promoter DNA methylation leading to uncontrolled cell growth. However, pharmacological demethylation can restore gene function and promote death of tumor cells. Thus, it is important to understand these relationships not only in individual genes, but also in combinations of genes that can then provide epigenetic signatures of various tumor types and lead to more rational treatments for diseases such as cancer. Oligonucleotide-based microarrays are new tools developed for high-throughput DNA methylation analysis that combine the most advanced chip technology and the bisulfite-treatment strategy. These arrays, coupled to informatics approaches, are likely to speed the identification of new genes that are silenced by hypermethylation in cancer and to classify the known genes into biologically and clinically relevant hierarchies. Such methods hold considerable promise for providing better prognostic information for the future of personalized medicine.

ACKNOWLEDGMENTS

The authors wish to thank Matthias Burger for the initial data analysis of non-Hodgkin's lymphomas and Kurt Berlin and Diane Peckham for the assistance in the preparation of this manuscript.

REFERENCES

Adorján P, Distler J, Lipscher E, Model F, Muller J, Pelet C, Braun A, Florl AR, Gutig D, Grabs G, Howe A, Kursar M, Lesche R, Leu E, Lewin A, Maier S, Muller V, Otto T, Scholz C, Schulz WA, Seifert HH, Schwope I, Ziebarth H, Berlin K, Piepenbrock C, Olek A. 2002. Tumor class prediction and discovery by microarray-based DNA methylation analysis. Nucleic Acids Res 30:e21.

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J, Jr., Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Staudt LM. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403:503–511.

Antequera F, Boyes J, Bird A. 1990. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. Cell 62:503–514.

Bird A. 1986. CpG-rich islands and the function of DNA methylation. Nature 321:209–213.

Bird A. 2002. DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21.

Clark SJ, Harrison J, Paul CL, Frommer M. 1994. High sensitivity mapping of methylated cytosines. Nucleic Acids Res 22:2990–2997.

Clarke PA, te Poele R, Wooster R, Workman P. 2001. Gene expression microarray analysis in cancer biology, pharmacology, and drug development: Progress and potential. Biochem Pharmacol 62:1311–1336.

Esteller M, Corn PG, Baylin SB, Herman JG. 2001. A gene hypermethylation profile of human cancer. Cancer Res 61:3225–3229.

Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. 1992. A genomic

sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci 89:1827–1831.

Gitan RS, Shi H, Yan PS, Huang TH-M. 2002. Methylation-specific oligonucleotide microarray: A new potential for high-throughput methylation analysis. Genome Res 12: 158–164.

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. 1999. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science 286: 531–537.

Jones PA, Baylin SB. 2002. The fundamental role of epigenetic events in cancer. Cancer Cell 3:415–428.

Jones PA, Laird PW. 1999. Cancer epigenetics comes of age. Nature Genet 21:163–167.

Lyon MF. 1999. Imprinting and X-chromosome inactivation. Results Probl Cell Differ 25:73–90.

Melki JR, Vincent PC, Clark SJ. 1999. Cancer-specific region of hypermethylation identified within the HIC1 putative tumour suppressor gene in acute myeloid leukaemia. Leukemia 13:877–883.

Model F, Adorján P, Olek A, Piepenbrock C. 2001. Feature selection for DNA methylation based cancer classification. Bioinformatics 17:S157–S164.

Rice JC, Massey-Brown KS, Futscher BW. 1998. Aberrant methylation of the BRCA1 CpG island promoter is associated with decreased BRCA1 mRNA in sporadic breast cancer cells. Oncogene 17:1807–1812.

Santini V, Kantarjian HM, Issa JP. 2001. Changes in DNA methylation in neoplasia: Pathophysiology and therapeutic implications. Ann Intl Med 134:573–586.

Smiraglia DJ, Rush LJ, Fruhwald MC, Dai Z, Held WA, Costello JF, Lang JC, Eng C, Li B, Wright FA, Caligiuri MA, Plass C. 2001. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. Hum Mol Genet 13:1413–1419.

Stirzaker C, Millar DS, Paul CL, Warnecke PM, Harrison J, Vincent PC, Frommer M, Clark SJ. 1997. Extensive DNA methylation spanning the Rb promoter in retinoblastoma tumors. Cancer Res 57:2229–2237.

Wei SH, Chen C-M, Strathdee G, Harnsomburana J, Shyu CR, Rahmatpanah F, Shi H, Ng SW, Yan PS, Nephew KP, Brown R, Huang TH-M. 2002. Methylation microarray analysis of late-stage ovarian carcinomas distinguishes progression-free survival in patients and identifies candidate epigenetic markers. Clin Cancer Res 8:2246–2252.

Widschwendter M, Jones PA. 2002. The potential prognostic, predictive, and therapeutic values of DNA methylation in cancer. Clin Cancer Res 8:17–21.

Yan P, Perry M, Laux D, Asare A, Caldwell C, Huang T-M. 2000. CpG island arrays: An application toward deciphering epgienetic signatures of breast cancer. Clin Cancer Res 6:1432–1438.

Yan PS, Chen C-M, Shi H, Rahmatpanah F, Wei SH, Caldwell CW, Huang TH-M. 2001. Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. Cancer Res 61:8375–8380.