2532–2534 *Nucleic Acids Research, 1997, Vol. 25, No. 12*

COBRA: a sensitive and quantitative DNA methylation assay

Zhenggang Xiong and Peter W. Laird*

Department of Surgery and Department of Biochemistry and Molecular Biology, University of Southern California School of Medicine, The Norris Comprehensive Cancer Center, Room 6415, Mailstop #73, 1441 Eastlake Avenue, Los Angeles, CA 90033, USA

Received March 25, 1997; Revised and Accepted April 18, 1997

ABSTRACT

We report here on a quantitative technique called COBRA to determine DNA methylation levels at specific gene loci in small amounts of genomic DNA. Restriction enzyme digestion is used to reveal methylation-dependent sequence differences in PCR products of sodium bisulfite-treated DNA as described previously. We show that methylation levels in the original DNA sample are represented by the relative amounts of digested and undigested PCR product in a linearly quantitative fashion across a wide spectrum of DNA methylation levels. In addition, we show that this technique can be reliably applied to DNA obtained from microdissected paraffin-embedded tissue samples. COBRA thus combines the powerful features of ease of use, quantitative accuracy, and compatibility with paraffin sections.

Recent reports of silencing of tumor-suppressor gene expression by DNA methylation (3) have emphasized the need for accurate, sensitive, reliable and quantitative methods to measure levels of DNA methylation at specific gene loci. Clinical and molecular epidemiological studies often involve the analysis of large numbers of paraffin-embedded tissue samples. Existing techniques to analyze DNA methylation levels have not been shown to satisfy the three criteria essential for such large-scale studies, which are: (i) quantitative accuracy, (ii) compatibility with paraffin sections, and (iii) applicability to large numbers of samples. We report here on a method that satisfies all these criteria. The technique relies on the recent innovation of Sadri and Hornsby (1), who showed that restriction digestion can be used to reveal DNA methylation-dependent sequence differences in PCR-amplified bisulfite-treated genomic DNA. We show that this feature can be exploited to determine DNA methylation levels at specific loci with a very high degree of quantitative accuracy. We show further that the technique can be applied to DNA samples obtained from microdissected paraffin-embedded sections. We refer to this procedure as COBRA (combined bisulfite restriction analysis).

The general outline of the method is depicted in Figure 1. Methylation-dependent sequence differences are introduced into the genomic DNA by the standard sodium bisulfite treatment and then PCR amplified (2). This combination of bisulfite treatment and PCR amplification results in the conversion of unmethylated cytosine residues to thymine and methylated cytosine residues to cytosine. This sequence conversion can lead to the methylation-dependent creation of new restriction enzyme sites (1) or it can lead to the methylation-dependent retention of pre-existing sites such as *Bst*UI (CGCG) (Fig. 1). The primers used in the PCR reaction do not contain CpG dinucleotides so that the amplification step does not discriminate between templates according to their original methylation status. Therefore, in the mixed population of resulting PCR fragments, the fraction that has a newly created or retained restriction site that contains a CpG(s) should be a direct reflection of the percentage DNA methylation at that site in the original genomic DNA. We have tested whether this feature can be used to quantitatively determine DNA methylation levels. We found that the most accurate and reliable method to quantitate the relative amounts of digested and undigested PCR products was to perform an unlabeled PCR reaction, followed by a purification step to ensure subsequent complete cutting, then restriction digestion, polyacrylamide gel electrophoresis, electroblotting, oligo hybridization and phosphorimager quantitation. This hybridization strategy allows flexibility provided by the choice of probe position, relative to the restriction sites. Results obtained with different probes can be used to corroborate DNA methylation values obtained for a particular CpG site.

We have used the human estrogen receptor (*ER*) gene as a model system to test our method. The human *ER* gene contains a CpG island at its 5′ end, which becomes increasingly methylated in colorectal mucosa with age and is heavily methylated in all human colorectal tumors analyzed (4). Figure 2 shows a COBRA analysis of the human colorectal tumor cell line HCT116 and of normal control DNA. The tumor cell line DNA shows extensive cutting by both *Bst*UI and *Taq*I, while the normal control DNA shows no cutting by *Bst*UI and minor cutting by *Taq*I. The calculated methylation levels are shown below the appropriate lanes. The restriction site for *Bst*UI will be retained only if both CpG dinucleotides within its recognition sequence are methylated. This is a more stringent requirement than for a site such as *Taq*I with a single CpG dinucleotide within its recognition sequence. Thus, the choice of restriction enzyme can be adjusted to accommodate different levels of DNA methylation in the samples to be analyzed.

*To whom correspondence should be addressed. Tel: +1 213 764 0650; Fax: +1 213 764 0158; Email: laird_p@froggy.hsc.usc.edu

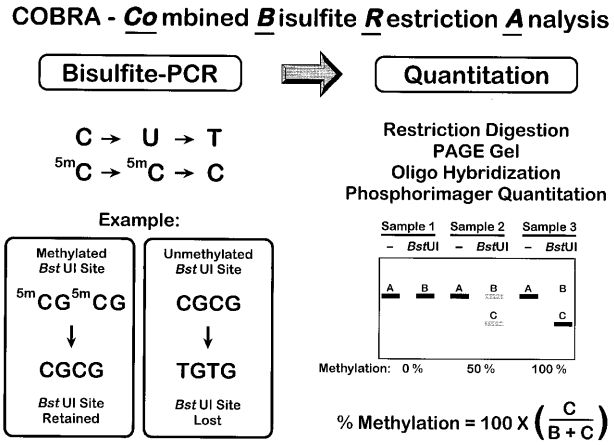

Figure 1. Outline of the COBRA procedure. COBRA consists of a standard sodium bisulfite PCR treatment followed by restriction digestion and quantitation step. During the first step, unmethylated cytosine residues are converted to thymine, whereas methylated cytosine residues are retained as cytosine. The consequences for the restriction site *Bst*UI are illustrated. The percentage of fully methylated *Bst*UI sites in a genomic DNA sample can be calculated from the ratio between the *Bst*UI-cleaved PCR product and the total amount of PCR product. The sodium bisulfite treatment and PCR reaction is performed essentially according to ref. 1. Approximately 1 µg of genomic DNA is used for the sodium bisulfite reaction. For samples with nanogram quantities of DNA (such as the DNA from paraffin-embedded sections), 1 µg salmon sperm DNA is added as carrier. The DNA is denatured in 0.2 M NaOH. Sodium bisulfite (Sigma, St Louis, MO) is then added to a final concentration of 3.1 M, and hydroquinone to a final concentration of 0.5 mM. The reaction is performed at 55°C for 16 h. The samples are desalted with the DNA Clean-Up System (Promega, Madison, WI). Subsequently, the DNA samples are desulfonated by 0.3 M NaOH and ethanol-precipitated. The PCR reactions are performed in a volume of 100 µl containing 1× Expand HF buffer (Boehringer Mannheim), 2.5 mM $MgCl_2$, 0.24 mM dNTPs, 1 µM of each primer and 2 U of Expand HF Enzyme Mix (Boehringer Mannheim). The PCR primers are designed to be complementary to the converted DNA sequences with no CpG dinucleotides in the corresponding region of original unconverted DNA. To ensure complete digestion of the PCR products in the subsequent restriction digestion step, the samples are first purified using 10K MWCO nanospin plus filters (Gelman Sciences) with two rinses of 200 µl water. The purified PCR products are then digested with a restriction enzyme with a recognition sequence containing CpG in the original unconverted DNA. Cleavage will occur only if the CpG sequence has been retained during the bisulfite conversion by a methylated status of the cytosine residue. It is essential to ensure that the bisulfite conversion is complete. Therefore, a control digest is performed with an enzyme such as *Hsp*92II, which has a recognition sequence (CATG) that should be destroyed by the bisulfite conversion. Any cleavage by *Hsp*92II would indicate either non-CpG DNA methylation or incomplete sodium bisulfite conversion. All restriction digests reactions are performed at the appropriate temperature for at least 4 h. The digested PCR products are separated on a 8% denaturing polyacrylamide gel and transferred to Zetabind charged membrane (American Bioanalytical) by electroblotting. The membranes are then hybridized with 5′-end-labeled oligonucleotides according to (8). Quantitation is performed with a Molecular Dynamics PhosphorImager.

We tested whether COBRA would provide quantitative results across a wide spectrum of DNA methylation levels by analyzing mixtures of the colorectal tumor cell line DNA and the normal control DNA in different ratios. The mixed samples were prepared prior to the bisulfite treatment and processed as independent samples throughout the entire COBRA procedure. The results are shown in Figure 2. Linear regression analysis of the data shows that the COBRA technique yields reliable quantitative results across a wide range of DNA methylation levels.

We next tested whether COBRA could be applied to the routine analysis of clinical tumor specimens that had been paraffin-embedded and microdissected (5). We analyzed seven pairs of

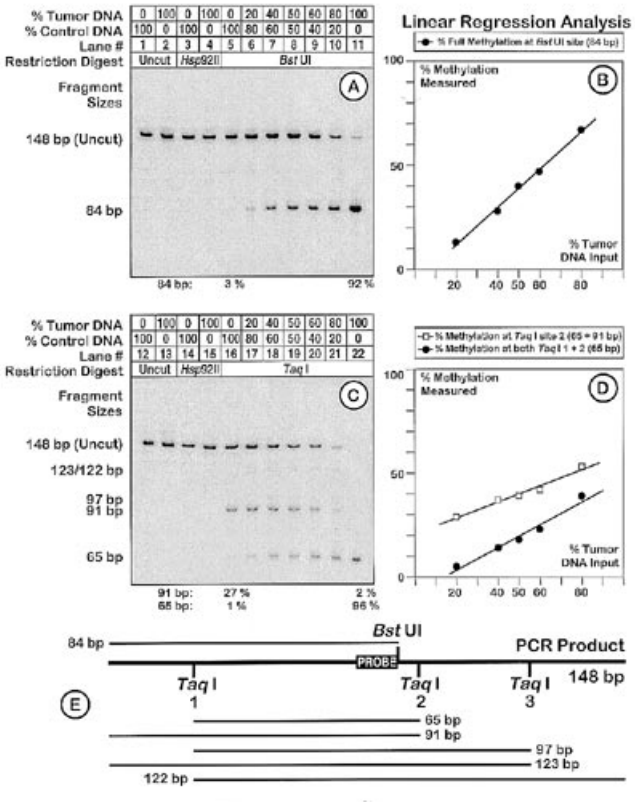

Figure 2. Determination of the quantitative accuracy of COBRA. Analysis was performed at the estrogen receptor CpG island on DNA samples derived from the human colorectal tumor cell line HCT116 [(**A**), lanes 2, 4 and 11]; [(**C**) lanes 13, 15 and 22]. Normal peripheral blood lymphocytes were used as a control [(A), lanes 1, 3 and 5]; [(C), lanes 12, 14 and 16]. (A) shows a COBRA assay with the restriction enzyme *Bst*UI, (C) with *Taq*I. The calculated methylation levels are shown below the appropriate lanes in (A) and (C). (**E**) Shows a map of the PCR product with the sizes of the predicted digestion products. There are three potential *Taq*I sites within the148 bp PCR product. The probe shown in (E) provides information on the levels of methylation at site #2. This value is obtained by calculating the combined signals of the 65 and the 91 bp bands. The signal of the 65 bp band alone represents the fraction of molecules in the original DNA sample with linked methylation at both *Taq*I sites #1 and #2. The levels of methylation of the other individual *Taq*I sites can be determined using probes at different locations in the PCR product. The tumor cell DNA and normal control DNAs were mixed by volume at various ratios as indicated (lanes 6–10 and 17–21) to determine the quantitative nature of the COBRA assay. Linear regression analyses of the results are shown in (**B**) and (**D**). The calculated DNA methylation percentages are plotted as a function of the percentage input tumor DNA. The correlation coefficient is 0.998 for the *Bst*UI digests (B). The correlation coefficient for the single cut at *Taq*I site #2 [(D), open squares] is 0.988. The correlation coefficient for linked methylation at both *Taq*I sites #1 and #2 [(D), closed circles] is 0.985. The values for the unmixed samples are not included in the linear regression analysis, since the DNA samples were mixed by volume, rather than by DNA amount. Experimental procedure: genomic DNA was isolated according to ref. 9 and processed as described in Figure 1. The PCR primers were designed to generate a 148-bp PCR product within the 5′ UTR of the human *ER* gene from the lower DNA strand of the bisulfite converted sequence (10). The 5′ primer OL-001 sequence (TCCTAAAACTACACTTACTCC) corresponds to positions +22 to +42 of the unconverted sequence (10). The 3′ primer OL-002 sequence (GGTTATTTG-GAAAAAGAGTATAG) corresponds to positions +169 to +147 (10). The probe OL-039 sequence (AAACCAAAACTC) corresponds to positions +92 to +103 in the unconverted sequence (10). *Hsp*92II digests (lanes 3, 4, 14 and 15) were performed to confirm that the sodium bisulfite conversion was complete (see Fig. 1).

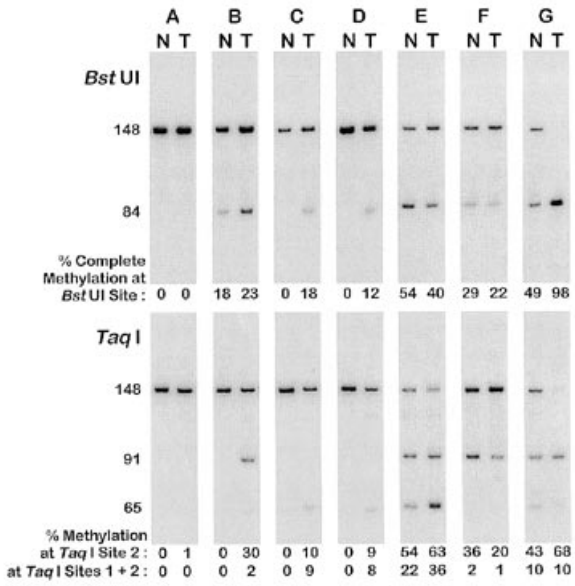

Figure 3. COBRA analysis of seven pairs (A–G) of microdissected human colorectal tumors and normal tissue margin. The tumor samples are indicated by a 'T'. The normal tissue margin is indicated by an 'N'. The calculated percentage of DNA methylation is shown below each lane. The values for the *Bst*UI digests reflect methylation of both CpG dinucleotides within the recognition site. Two values are given for the *Taq*I digests. The upper value represents the percentage of signal in both the 65 and the 91 bp bands added together. This value corresponds to the percentage methylation at *Taq*I site #2 (see map in Fig. 2). The lower value represents the percentage of signal in the 65 bp band, which reflects the fraction of original DNA molecules with linked methylation at both *Taq*I sites #1 and #2. Experimental procedure: DNA was isolated from microdissected samples as described by Shibata (5). Salmon sperm DNA (1 µg) was added as a carrier to each sample before the bisulfite treatment.

human colorectal tumors and normal adjacent tissue margins for methylation status of the *ER* CpG island. The results are shown in Figure 3. Issa *et al*. had reported complete methylation of a *Not*I site in all human colorectal tumor samples analyzed (4). Using COBRA, we see a much more diverse spectrum of DNA methylation levels, from a complete lack of detectable DNA methylation in patient A to very high levels of DNA methylation in the tumor sample of patient G. These results confirm that the human *ER* CpG island is hypermethylated in most colorectal tumors, but suggest that more detailed information on methylation levels could be obtained using COBRA. These results also show that the COBRA method is well suited for the analysis of large numbers of paraffin-embedded clinical specimens.

Southern blot analysis is a fairly quantitative method to analyze DNA methylation levels at specific gene loci, but it is incompatible with DNA isolated from paraffin sections. Methylation-sensitive restriction enzyme digestion followed by PCR is prone to false-positive results since even low levels of spurious incomplete digestion can result in a PCR product. This problem is exacerbated in samples derived from paraffin sections. COBRA circumvents this problem by restriction digestion of a purified PCR product, rather than of the original genomic DNA. A recently described bisulfite-PCR-based method called MSP (6) is an excellent alternative, but tends to be a more qualitative, rather than quantitatively accurate method. COBRA and the Ms-SNuPE method described in the adjacent paper (7) are the first techniques to combine the attractive features of rapid ease of use, highly quantitative results and compatibility with paraffin-embedded samples.

ACKNOWLEDGEMENTS

We thank Dr Darryl Shibata (Department of Pathology, USC) and Dr Anna Wu (Department of Preventive Medicine) for DNA samples of microdissected tissues. We thank Dr Jeffrey Weber (Department of Medicine, USC) for a sample of normal human peripheral mononuclear cells. This work was supported by grant #IRG-21-36 from the American Cancer Society and grant 2CA-14089-21 from the National Cancer Institute. P.W.L. is the 1996 recipient of the Stop Cancer Lily Opas Research Career Development Award.

REFERENCES

1 Sadri,R. and Hornsby,P.J. (1996) *Nucleic Acids Res.*, **24**, 5058–5059.
2 Frommer,M., McDonald,L.E., Millar,D.S., Collis,C.M., Watt,F., Grigg,G.W., Molloy,P.L. and Paul,C.L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1827–1831.
3 Jones,P.A. (1996) *Cancer Res.*, **56**, 2463–2467.
4 Issa,J.P., Ottaviano,Y.L., Celano,P., Hamilton,S.R., Davidson,N.E. and Baylin,S.B. (1994) *Nature Genet.*, **7**, 536–540.
5 Shibata,D. (1992) In Herrington,C.S. and McGee,J.O.D. (eds), *Diagnostic Molecular Pathology: A Practical Approach*. Volume II. IRL Press, Oxford.
6 Herman,J.G., Graff,J.R., Myöhänen,S., Nelkin,B.D. and Baylin,S.B. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 9821–9826.
7 Gonzalgo,M.L. and Jones,P.A. (1997) *Nucleic Acids Res.*, **25**, 2529–2531.
8 Church,G.M. and Gilbert,W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
9 Laird,P.W., Zijderveld,A., Linders,K., Rudnicki,M.A., Jaenisch,R. and Berns,A. (1991) *Nucleic Acids Res.*, **19**, 4293.
10 Green,S., Walter,P., Kumar,V., Krust,A., Bornert,J.-M., Argos,P. and Chambon,P. (1986) *Nature*, **320**, 134–139.