

Rapid quantification of DNA methylation by measuring relative peak heights in direct bisulfite-PCR sequencing traces

Minghong Jiang^{1,2}, Yuhao Zhang¹, Jing Fei¹, Xinxia Chang³, Weiwei Fan¹, Xueqing Qian¹, Tianbao Zhang⁴ and Daru Lu¹

Various technologies are currently available to quantify DNA methylation. However, rapid and simple methods for determining the DNA methylation status of CpG sites in genes still remain elusive. In this report, we describe a novel method for the rapid quantification of CpG methylation on the basis of direct bisulfite-PCR sequencing method. According to the principles of bisulfite-PCR, converting unmethylated cytosines to thymine while leaving methylated cytosines unchanged, we regard the CpG site as a SNP and estimate the methylation status of cytosines in the given CG dinucleotides by measuring the ratio of the cytosine peak height to the sum of cytosine and thymine peak heights in automated DNA sequencing traces. Furthermore, we take several effective measures to break through the 'bottleneck' problems that render the routine bisulfite sequencing method unsuitable for quantitative methylation. In comparison with pyrosequencing and bisulfite-cloning sequencing, our method is confirmed to be a simple, high-throughput and cost-effective technology for determining the methylation status of specific genes. Accordingly, this novel method is anticipated to be an efficient and economical alternative tool for rapid quantification of methylation patterns in screening large numbers of clinical samples across multiple genes.

Laboratory Investigation (2010) **90**, 282–290; doi:10.1038/labinvest.2009.132; published online 14 December 2009

KEYWORDS: DNA methylation; bisulfite sequencing; PCR; CpG sites; pyrosequencing

DNA methylation at the carbon-5 position of the cytosine base in CpG dinucleotides is a common chemical modification of DNA in eukaryotes. It has been shown to have an important role in the transcriptional regulation of multiple physiological processes. Normal DNA methylation changes are involved in embryogenesis, cell differentiation and aging,^{1,2} whereas aberrant DNA methylation has a close connection with several chronic diseases, such as cancer and diabetes.^{3–5} Thus, the simple and rapid detection of aberrant DNA methylation is generating more and more interest. It not only enhances our understanding of how DNA methylation is regulated but is also helpful to investigations into the relationship between methylation and disease development.^{6,7}

Recently, many approaches for quantitative assessment of DNA methylation depending on sodium bisulfite treatment

of genomic DNA and subsequent PCR amplification have been developed and widely used.^{8–10} In these methods, when DNA is treated with a high concentration of bisulfite, unmethylated cytosine is converted into uracil, whereas methylated cytosine remains intact. If a target gene is amplified by PCR, the PCR product will contain thymine at a general cytosine location and cytosine at a methyl cytosine location. On the basis of this principle, several methods are used to analyze the methylation status of a target gene with high sensitivity and specificity. Standard methylation-specific PCR (MSP) is one of the simplest methods, but is not quantitative. Several quantitative variations on the basis of MSP, such as MethyLight,¹¹ HeavyMethyl¹² and MethylQuant,¹³ improved DNA methylation quantification in the CpG dinucleotide. However, these strategies can only determine the methylation

¹State Key Laboratory of Genetic Engineering, School of Life Science and Institute of Biomedical Sciences, Fudan University, Shanghai, China; ²National Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Peking Union Medical College, Tsinghua University and Chinese Academy of Medical Sciences, Beijing, China; ³Department of Endocrinology, Zhongshan Hospital, Fudan University, Shanghai, China and ⁴Department of Health Toxicology, College of Basic Medical Sciences, Second Military Medical University, Shanghai, China

Correspondence: Dr DR Lu, PhD, State Key Laboratory of Genetic Engineering, Fudan University, School of Life Science and Institute of Biomedical Sciences, 220 Handan Road, Shanghai 200433, China.

E-mail: drlu@fudan.edu.cn

Received 8 May 2009; revised 19 August 2009; accepted 29 September 2009

level for one or two CpGs overlapped by PCR primers, leaving other sites unexplored. As such, their application and quantitative utility is restricted. Analysis of all CpG dinucleotides within a given sequence will provide a broader view of DNA methylation levels. Many studies have already shown that very few or even single CpG methylation changes are sufficient to epigenetically alter the expression of a gene through quantitative methylation analysis of every CpG site located in its promoter.^{14–17} At present, the most popular methods for quantification of every CpG methylation event occurring in a specific region involve pyrosequencing and DNA sequencing. Pyrosequencing, in an unbiased manner, is a newly emerging method, although its intrinsic short-read sequencing (normally only up to 30 bp at a time) presents a disadvantage in comparison with DNA sequencing.^{6,18} As for sequencing-based DNA methylation techniques, the cloning-based sequencing protocol is one of the most widely accepted strategies.^{19,20} However, this approach has a shortcoming: multiple laborious steps (such as bisulfite treatment, cloning of PCR fragments, construction of recombinant vectors, identification of positive clones and DNA sequencing). This technique requires cloning of the PCR product before sequencing for adequate sensitivity, and therefore is a very labor-intensive and time-consuming method unsuitable for high-throughput sample analysis. Usually, ~10 clones must be examined to determine the degree of methylation from a mixture of PCR fragments of a sample, inevitably reducing the statistical power of the sequencing data. A strategy that forgoes cloning-based sequencing and simply directly sequences PCR products will become an ideal high-throughput platform. At present, the approach of direct sequencing has failed to gain acceptance in the field of epigenetic study as a reliable method for quantification of methylation from sequencing chromatograms of poor data, because of high background noise and overscaled cytosine signals.^{20,21} On the basis of direct bisulfite sequencing, Paul and Clark²⁰ and Lewin *et al*²² have independently developed a modified approach using an innovative protocol and novel algorithm to rectify distortions in sequencing trace files. Despite the richness of information obtained from their sequencing methods, these two methods are too laborious, time-consuming and expensive to use in most clinical laboratories. Therefore, simple, rapid and inexpensive alternative methods, which allow accurate quantitative assessment of DNA methylation especially for clinical samples, are desired based on the technology of bisulfite genomic sequencing.

In this study, we introduce a novel method for evaluating DNA methylation status using peak height information obtained from four-dye trace files in direct sequencing of bisulfite-treated PCR products. In our previous studies, we quantified methylation levels of the glucokinase (*Gck*) promoter in rats of various ages using this new strategy and the classical cloning-based sequencing of PCR products, and found that results of the two methods share high similarity.²³ To assess the feasibility and reliability of this approach in

detail, we examined it in parallel with the cloning-based sequencing of PCR products from the bisulfite-treated *Gck* promoter in rat liver tissues and BRL line. In addition, to examine linearity and precision, we compared the method with the pyrosequencing assay using a dilution matrix of known concentrations of mixed methylated and unmethylated *Gck* promoter regions.

MATERIALS AND METHODS

Cell Lines

The normal rat liver cell line BRL was obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). BRL cells were cultured in high-glucose Dulbecco's modified Eagle's medium (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum (Gibco, USA).

Animals

Male, 3-week-old Wistar rats (Animal Developmental Centre, Chinese Academy of Sciences) were given free access to commercial laboratory animal chow. Experiments were carried out after 16 h of fasting. Experimental procedures involving the use of animals were conducted in accordance with the NIH Guidelines and were reviewed and approved by the Animal Use and Care Committee of Fudan University.

DNA Isolation and Bisulfite Sequencing PCR Amplification

Genomic DNA obtained from rat liver tissues and BRL cells were isolated by digestion with proteinase K and phenol–chloroform extraction as described previously.¹⁹ DNA was purified again using Nucleic Acid Purification Kits MagExtractor Genome (Toyobo, Osaka, Japan) following the manufacturer's instructions.

Bisulfite modification of genomic DNA was carried out according to previously reported methods.²³ In general, 2 µg of genomic DNA was treated with sodium bisulfite and extracted to a final volume of 40 µl of modified genomic DNA. The basal promoter region of the rat hepatic *Gck* gene, spanning a 601-nucleotide (nt) fragment with 11 CpG sites from nt –518 to nt +83, was amplified with nested primers (Table 1) under amplification conditions as described previously.^{23,24}

A schematic diagram of primers and CpG dinucleotide positions within the bisulfite-converted *Gck* gene promoter sequence (all C converted to T, except for CpG sites.) is shown in Figure 1. The first cycle reaction mixture (10 µl) contained 2 µl of modified DNA template and 2 × TaqPCRMaster (Tiangen Biotech, Shanghai, China). The first PCR products (2 µl) were subjected to a second round of PCR (10 µl). Liver-type pyruvate kinase (*LPK*) promoter regions essential for basal transcriptional activity occurred between nt –280 and nt –191 and included 14 CpG sites.²⁵ The 471-bp region was amplified with two rounds of PCR using primers designed using the MethPrimer software (<http://www.ucsf.edu/urogene/methprimer/index1.html>)

Table 1 Primer sequences for bisulfite sequencing PCR of hepatic *Gck*, *L-PK* and *Glut2* and pyrosequencing of *Gck* promoter

Gene	Forward primer (5' → 3')	Reverse primer (5' → 3')
<i>Bisulfite sequencing PCR (BSP)</i>		
<i>Gck</i> 1st	GAATTTTATAGAAGAGTTTGAATGTTTTGG	CACACCTTATAATATCCATAACCATCTC
<i>Gck</i> 2nd	GGGTGTAGGGTAGTTAGAGGATTTG	CCTAACTCCTAAAACACCTATTAC
<i>LPK</i>	TTTTGTGGGTAGGATGTTAATATAG	TTCCAACTCCAACAATTCTCTAAC
<i>Glut2</i>	GGGGATTGTGAAGATTGATAAGTAG	TTCTCAAAAACACTCAAAAATTCCTA
<i>Pyrosequencing of Gck</i>		
Pyr	TATGGTGAAGTTTATGATTGTGTGA	Biotin-AAATAACTAAAACCTACCCACTCCC
PyrS	TGAAGTTTATGATTGTGTGA	

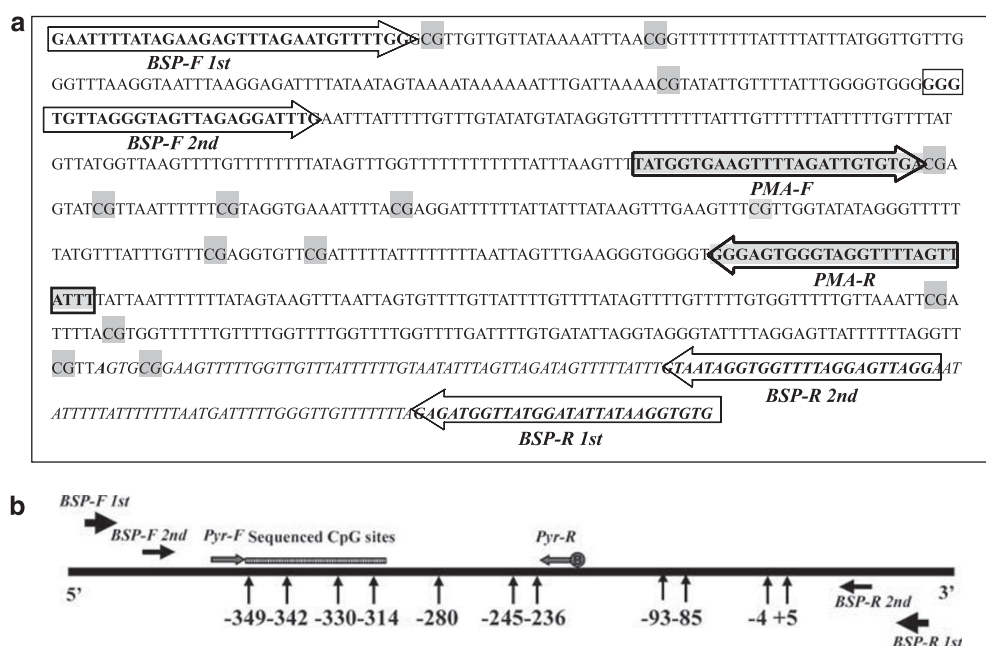


Figure 1 Diagrammatic representation of primer alignment and CpG dinucleotide positions on the *Gck* promoter region. (a) Bisulfite-treated *Gck* promoter sequence (all C converted to T, except for CpGs). CpG dinucleotides in gray boxes indicate nucleotides detected by BSP and pyrosequencing assays. (b) Schematic diagram of primers (indicated by arrows) and 11 CpG sites (indicated by up arrows) from nt -518 to nt +83 within the *Gck* promoter. BSP-F, bisulfite sequencing PCR forward primers; BSP-R, bisulfite sequencing PCR reverse primers; PMA-F, pyrosequencing methylation assay forward primers; PMA-R, pyrosequencing methylation assay reverse primers.

(Table 1). Bisulfite sequencing PCR amplification (BSP) of the *LPK* promoter was performed using our novel program (Table 2). In brief, the first round of the PCR reaction (9 μ l) contained 0.7 μ l of modified genomic DNA and was overlaid with mineral oil to form a vapor barrier. The cycling conditions consisted of an initial denaturation at 96°C for 5 min, at which point 0.5 μ l reverse primer (10 μ M) was added to the PCR mixture, followed by 2 cycles of 96°C for 1 min, 60°C for 2 min and 72°C for 2 min. When the temperature again increased to 96°C, which was the beginning of the next phase of 8 cycles (96°C for 1 min, 60°C for 2 min and 72°C for 2 min), 0.5 μ l of the forward primer was added to the

PCR mixture. Thereafter, 30 cycles of 96°C for 30 s, 53°C for 45 s and 72°C for 45 s were performed, followed by a final extension of 72°C for 7 min. The first-round products (0.5 μ l) were subjected to a second round of PCR under the following conditions: for 1 min at 96°C, for 2 min at 58°C and for 1 min 30 s at 72°C for 35 cycles. Another amplified region, glucose transporter type 2 (*Glut2*) promoter, used this optimization program with a slight modification as described in Supplementary material.

After amplification, the size and quality of PCR products (2 μ l) were visualized on 2% agarose gel. For direct sequencing, another 2 μ l of PCR products was purified with

Table 2 Optimization program used for BSP amplification

First amplification			Second amplification		
Cycles	Temperature (°C)	Time	Cycles	Temperature (°C)	Time
2	96	5 min	40	96	5 min
	96	1 min		96	1 min
	60 (64–56)	2 min		58 (62–55)	45 s
	72	2 min		72	45 s
8	96	1 min		72	7 min
	60 (64–56)	1 min			
	72	1 min			
	96	1 min			
30	53 (58–50)	45 s			
	72	45 s			
	72	7 min			

The original reaction mixture (9 μ l) for the first-round PCR contains no primers. When the temperature is increased to 96°C, the thermal cycler is paused at that temperature, the closed tubes are opened, and 0.5 μ l of 10 μ M reverse primer and 0.5 μ l of 10 μ M forward primer are added to the PCR mixture (under the mineral oil) in different denaturation steps. The tubes were then closed to continue the cycle.

exonuclease I (ExoI) and shrimp alkaline phosphatase (SAP) enzymes (United States Biochemical, Cleveland, OH, USA) to eliminate unincorporated dNTPs and primers. Enzymatic purification was carried out in a 7 μ l PCR mix by adding 3 Units of ExoI, 1 Unit SAP and 10 \times SAP reaction buffer, which were incubated for 60 min at 37°C, followed by 15 min at 80°C for enzyme denaturation. The sequencing reaction was carried out using 1 μ l ExoI/SAP-purified PCR amplifications, 0.5 μ l Big Dye Terminator Kit (Applied Biosystems, Foster City, CA, USA) and 1 μ l (3.2 μ M) of the reverse primer from the second-round BSP in a total volume of 5 μ l. The sequencing reaction was then performed as follows: 96°C for 1 min and 25 cycles of 96°C for 10 s, 50°C for 5 s and 60°C for 4 min. Finally, the sequencing reaction fragments were purified using an EDTA/ethanol protocol and sequenced on an ABI prism 3730 Genetic Analyzer (Applied Biosystems) with Dye terminators (Perkin-Elmer, Foster City, CA, USA).

Methylation Quantification of DNA Sequencing Data

The percentage of methylation was calculated as the peak height of C *vs* the peak height of C plus the peak height of T for each CpG site as shown in the computer-generated sequencing chromatogram extracted from the Chromas program (Version 2.32, Technelysium) (Supplementary Figure 1). A single C at the corresponding CpG site was considered as 100% methylation, a single T as no methylation and overlapping C and T as partial methylation (0–100%).

Cloning and Sequencing

PCR products were purified using a PCR purification kit (Qiagen, Hilden, Germany) and were cloned into the pGEM-T vector (Promega, Madison, WI, USA) according to the manufacturer's protocol. A total of 10 positive clones per sample were isolated using a kit (QiaPrep Spin Plasmid Miniprep; Qiagen) and sequenced. The methylation level for each CpG site was calculated by dividing the number of methylated CpGs in each site by the total number of clones sequenced.

Generation of DNA Methylation Standards

To further assess the reliability of BSP for quantification of methylation levels, we prepared control DNA standards containing various methylation percentages. Two subclones, one methylated and the other unmethylated, at all CpG sites in *Gck* promoter fragments of pGEM-T vectors, were selected from the bisulfite clones of rat liver tissues (as described above) and amplified with *Gck* pyrosequencing primers, Pyr-F and Pyr-R (Table 1), which were designed near the CpG-dense sites within the 601-bp fragment. The new PCR product (201 bp) contained seven CpG sites (Figure 1). The PCR reaction was carried out at 95°C for 10 min, followed by 30 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 1 min. The amplified methylated (M) and unmethylated (U) PCR products were then adjusted to a concentration of 50 ng/ μ l, mixed in proportions (M:U) of 0:6, 1:5, 2:4, 3:3, 4:2, 5:1, 6:0, and yielded samples with the following methylation levels: 0, 16.7, 33.3, 50, 66.7, 83.3 and 100%, respectively. Each DNA mixture was diluted to a final volume of 20 μ l with ddH₂O and was sequenced as described above and pyrosequenced simultaneously.

Pyrosequencing Assay

Single-stranded DNA from 8 μ l of each PCR mixture sample was purified according to the PSQ 96-sample preparation guide using a vacuum filtration sample device as per the manufacturer's directions (Biotage, Charlottesville, VA, USA). The single-stranded product was annealed to 0.3 μ M concentration of the sequencing primer Pyrs, placed at 85°C for 2 min and thereafter cooled to room temperature. Pyrosequencing was then performed on a PSQ HS 96 system using the Biotage reagent kit (Biotage) according to the manufacturer's instructions. Raw data were analyzed using the methylation quantification algorithm of the provided software.

Statistical Analysis

Data were given as mean \pm s.d. Linear regression, s.d. and evaluation of experimental data were performed using Microsoft Excel software. Statistical analysis was performed using Student's *t*-test. Significance was defined as $P < 0.05$.

RESULTS

Direct BSP Amplification and Quantification of DNA Methylation

DNA collected from all samples was treated with bisulfite, and then the region of interest was amplified by PCR with optimized PCR programs for increasing the specificity and yield of PCR products. This process converts the originally unmethylated CpG dinucleotides to TpG, while conserving the originally methylated CpGs. When the PCR products were visualized, unique bands of desired sizes were treated with ExoI/SAP and other procedures before direct sequencing.

In this study, we amplified *Gck*, *LPK* and *Glut2* promoters from bisulfite-treated genomic DNA of rat liver tissues and BRL cells. The specific PCR bands are shown on 2.0% agarose gel (Supplementary Figure 2A–C). As sequences were read with a reverse primer, the inversed/complemented files made it convenient to read potential CpGs. Representative sequencing chromatograms of *Gck*, *LPK* and *Glut2* (Figure 2a and c, Supplementary Figure 3) indicated high-quality data with sharp or evenly distributed peaks, and almost no background noise was observed. In addition, all sites of non-CpG cytosines were displayed as thymines in these figures, which implied that the bisulfite-induced conversion of unmethylated cytosines to uracil was complete in our DNA samples, and this serves as the basis for the subsequent calculation of methylation rates by measuring relative peak heights of cytosine (C) and thymine (T) peaks at each CpG in the sequencing traces. To examine the feasibility of our strategy, we assessed the methylation status of 11 CpG sites located in the promoter of the *Gck* gene (Figure 1).

The *Gck* promoter was amplified from rat liver tissues ($n=5$) and BRL cells ($n=3$), and the percentage of methylation at every CpG site was calculated as $C/(C+T)$. As shown in Figure 2, the methylation degree of rat liver samples was relatively even (45–95%) at 11 CpG sites (Figure 2b), whereas the level of BRL cells ranged from 10 to 100% methylation (Figure 2d).

Comparison Between Direct BSP Sequencing and Cloning-Based Sequencing for the Quantification of DNA Methylation

To further validate our findings, we compared these direct sequencing outcomes with those of sequencing from cloned PCR products obtained from the same samples. We sequenced 10 clones of each individual PCR reaction. As shown in Figure 2b and d, methylation states of individual CpG sites determined by cloning-based sequencing were similar to those measured by the direct BSP sequencing method. In addition, there were no significant differences ($P>0.05$) in methylation levels between the two methods. We also compared the two methods for quantification of DNA methylation in the hepatic *Gck* promoter of three different age groups of rats from a previous study.²³ Our method indicates a very good overall correlation with cloning-based sequencing. The results in Figure 2 also showed high consistency between the

two methods, except for several analyzed sites of the *Gck* promoter. However, we failed to identify which approach caused the differences. The number of clones might not be fully representative of the sequence in the template mixture; thus, a more precise estimation of direct BSP sequencing is required to allow comparison of quantification data with the direct sequencing method.

Accuracy Determination of BSP Sequencing for DNA Methylation Quantification

To more precisely estimate our the sensitivity and linearity of our method, various samples of different methylation rates from mixtures of fully methylated and non-methylated *Gck* promoter fragments amplified from two clones (electrophoresis in Supplementary Figure 2D) at different proportions (representing 0, 16.7, 33.3, 50, 66.7, 83.3 and 100% methylation) were detected by direct sequencing and pyrosequencing assays, respectively. Figure 3a showed typical pyrosequencing programs of these samples. Unmethylated DNA was calculated automatically as <4% methylation. With the expected methylation rates gradually increasing, the value calculated by $C/(C+T)$ increased proportionally from 0 to 100% methylation. A relatively good linear relationship ($R^2=0.9739$) was observed in Figure 4a.

Similarly, PCR products from the same mixed samples in the pyrosequencing assay were measured by our method. As shown in Figure 3b, the representative DNA sequence chromatograms showed only a T peak, whereas completely methylated sites appeared as only a C peak at the CpG sites in the chromatograms. The ratio of C gradually increased with the expected methylation rates. Figure 4b showed a good correlation ($R^2=0.9673$) with the data obtained from our system. This indicated that DNA methylation and peak height ratios in these standard samples increased proportionally, allowing the determination of different methylation states for the test samples. In addition, our method showed smaller system deviation than did the pyrosequencing method (slope of 0.9297 vs 0.8796). However, the standard curve did not cross the origin (intercept value of 4.5), which implied a <5% background noise in the observed samples, which might result from failing to distinguish between a real peak and background noise when a small C (or T) peak existed in a sequencing trace.

In addition, to evaluate the interassay variability of the test, we detected the PCR mixture amplification from the *LPK* promoter of different methylation rates, using the same procedure as described above. That standard curve (Supplementary Figure 4) showed a similar correlation and system deviation with Figure 4b, which implied that our method could accurately quantify DNA methylation levels on multiple independent CpG sites in different targets sequence.

DISCUSSION

DNA methylation is a crucial epigenetic modification of the genome that is involved in regulating many human diseases.

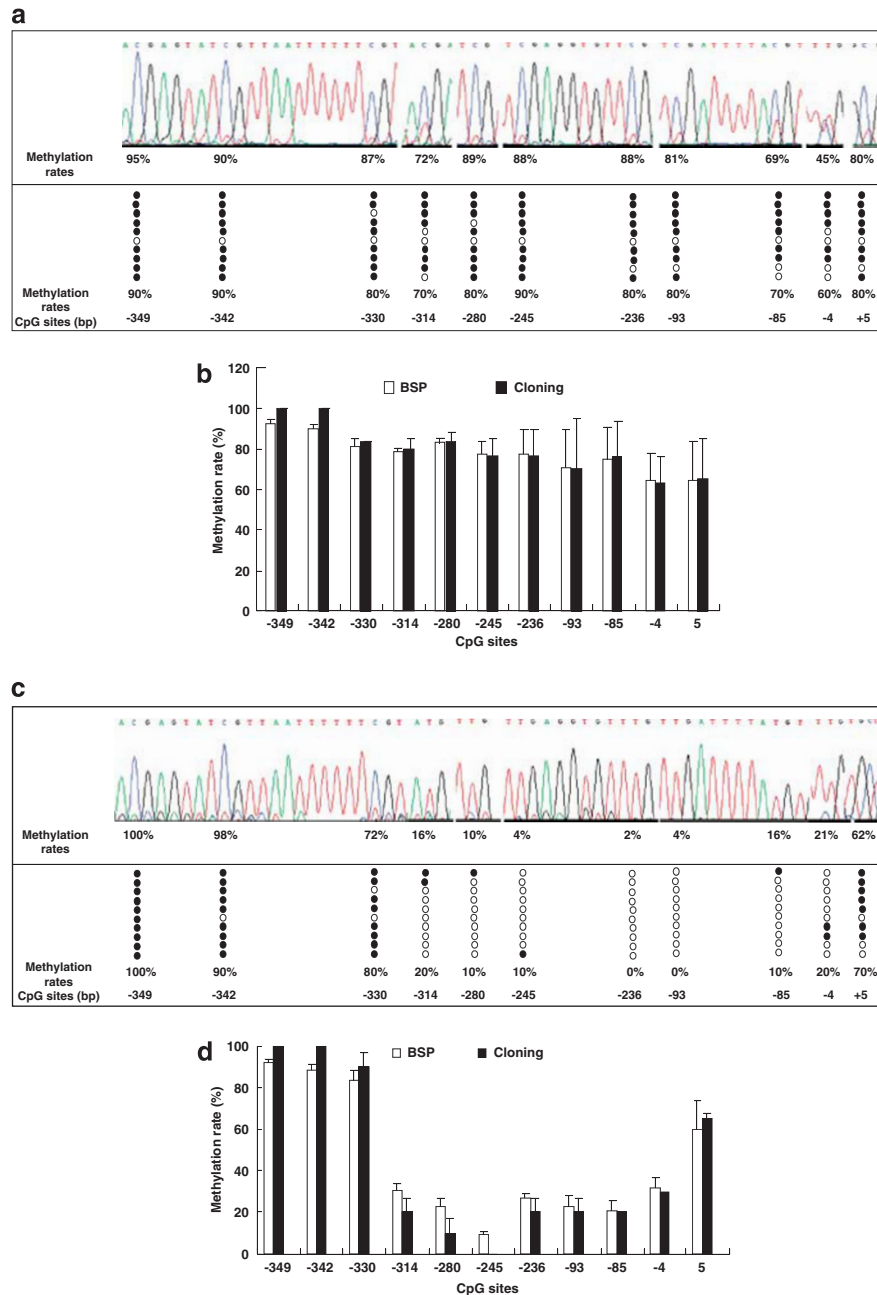


Figure 2 Validation of the present method with cloning-based sequencing of the *Gck* promoter amplified from DNA isolated from rat livers and BRL cells. A representative comparative result of methylation states between direct bisulfite-PCR sequencing chromatograms (upper) and cloning-based sequencing (lower) of the same fragment from a rat liver (**a**) and the BRL cell line (**c**). Open circles show no methylation, and closed circles indicate methylated cytosine. Comparison of methylation levels of the *Gck* promoter from rat livers ($n=5$) (**b**) and the BRL cells ($n=3$) (**d**) using the two methods.

With the biological and clinical importance of DNA methylation increasing rapidly, the technology for analyzing DNA methylation patterns has undergone swift development in recent years.

This report describes a simple and rapid DNA methylation quantification method, which consists of treating genomic DNA extracted from a sample with bisulfite, amplifying a target gene by PCR, purifying the BSP products to remove dNTP and primers, PCR sequencing and finally measuring

the content of methyl cytosine on the basis of the trace file data generated from sequencing chromatograms. Compared with conventional bisulfite sequencing, this method does not include complicated procedures of cloning PCR products into individual colonies and sequencing each clone. We made a special effort to optimize the protocols of bisulfite conversion and PCR amplification to render them applicable for direct DNA methylation analysis from bisulfite-PCR sequencing. We first improved measures for bisulfite

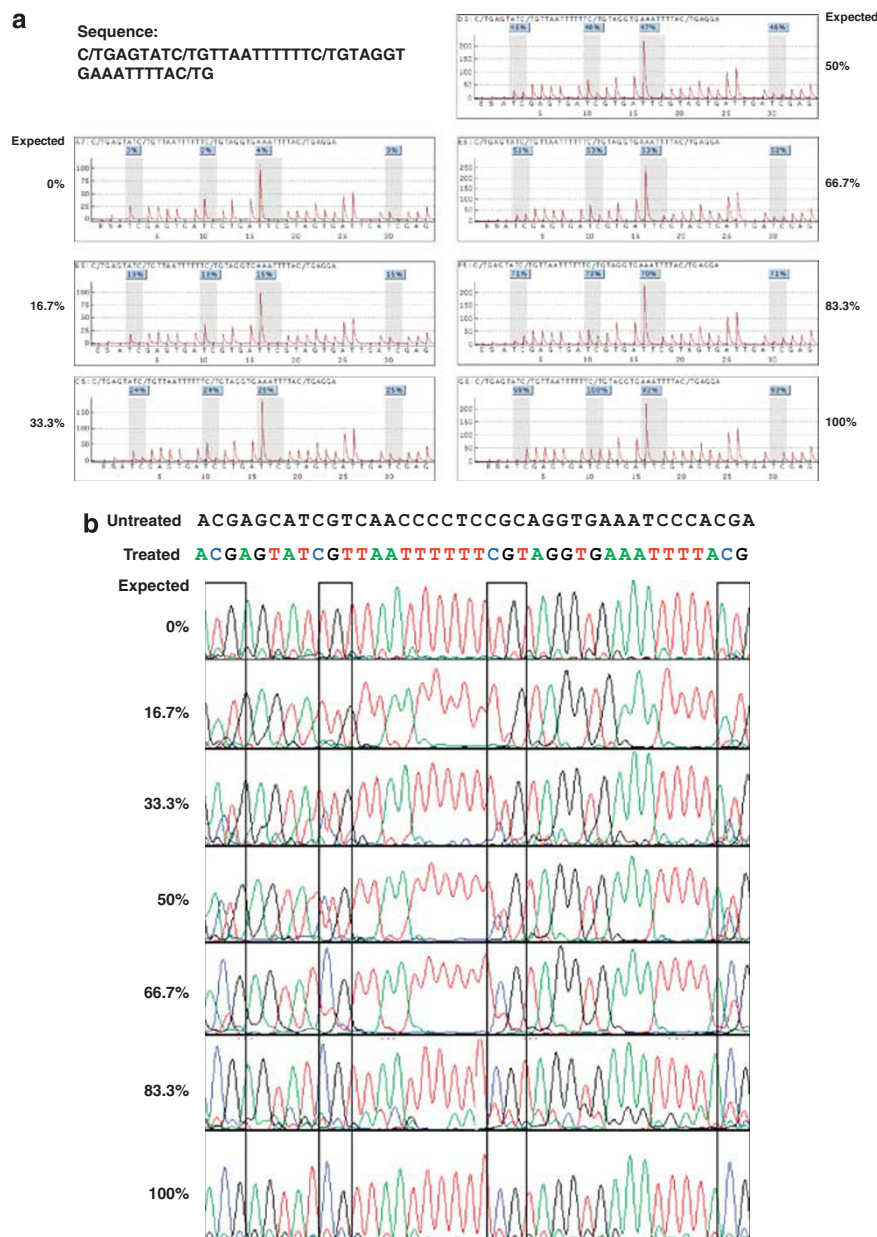


Figure 3 Representative pyrograms and sequencing chromatograms for different methylation levels of the *Gck* promoter region by pyrosequencing and our method. **(a)** Pyrosequencing assay. The expected and experimental values of methylation levels are shown at the two sides of each pyrogram or on the top of each CpG site. Shaded bars encompass C/Ct pairs. **(b)** Direct BSP sequencing. The expected values are shown on the left side of each chromatogram. The open boxes indicate the CpG sites.

treatment of purified genomic DNA to ensure the efficiency of bisulfite conversion. In addition, to increase the specificity and efficiency of PCR products, we invented a set of PCR for various special gene BSP amplifications (Table 2). First, we designed primers for bisulfite-converted DNA using the Methprimer software. In general, primer selection is one of the most critical steps in methylation analysis; therefore, it would be better to design two or three pairs of primers and then choose the most specific one or two from them to amplify target genes using our novel PCR program. PCR was performed in three cycle programs with two different

annealing temperatures. In the first cycle program, one of the primers (usually the reverse primer) was added in the denaturation phase and amplified for two cycles with annealing temperatures at $\sim 60^{\circ}\text{C}$ ($2\text{--}4^{\circ}\text{C}$ higher than the T_m of most primers with a length of ~ 20 bp) to yield perfect DNA–DNA matches at a sufficiently high temperature with one primer in a one-directional reaction. Subsequently, in the next denaturation phase of the second cycle program, the other primer (usually the forward primer) was added and amplified for eight cycles with the same annealing temperatures to produce high-specific PCR products. This was then followed by the

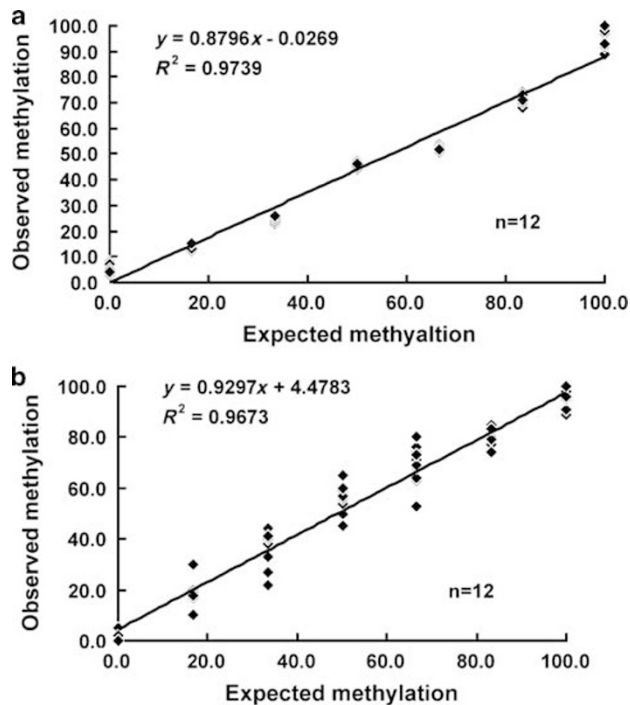


Figure 4 Linear relationship between observed and expected methylation levels by the pyrosequencing assay and our method. Quantitative results from multiple mixing experiments were compiled for both methods. The average quantification value for the same M/U mixture from 3 repeated sequences (a sequenced DNA fragment spanning 4 CpG sites, a total of 12 CpG sites) was plotted against its expected value. Linear regressions are shown for both pyrosequencing (a) and our method (b).

third cycle program in which the annealing temperature was set at 53°C (2–4°C lower than T_m of most primers with a length of ~20 bp) with 30 cycles to produce sufficient PCR products to the fullest extent. If there are nonspecific bands or weak products displayed in a gel, we can change those annealing temperatures by 2–5°C or continue into the second round of PCR, which is a typical PCR reaction with a 58°C annealing temperature and can also be modified by 2–5°C using nested primers, semi-nested primers or even the same primer pair as the first round of PCR.

Using the above-mentioned improved multiplex reactions, we have already successfully amplified various high-specificity BSP products, such as *LPK* promoters (Supplemental Figure 2A), *Glut2* promoters (Supplementary Figure 2B), *Gck* promoters (Supplementary Figure 2C) and other DNA sequences of interest (Supplementary Figure 2E) from various samples.

In general, with the obtained PCR products for methylation analysis, the traditional method is to clone and sequence at least 10 individual clones. However, this is a time-consuming work. Considering the fact that bisulfite treatment can convert unmethylated cytosines to thymines or leave methyl cytosine unchanged in PCR products, each cytosine within the CpG dinucleotide can obtain two peaks at one site (C/T) and can be regarded as a SNP. Although the direct

quantification of SNPs by measuring peak height ratio from direct PCR sequencing data is regarded as an accurate and sensitive method,²⁶ the quantification of cytosine methylation by direct BSP sequencing is considered impossible because it faces several challenges, such as poor signal quality, overscaled cytosine signals and base-caller artifacts,^{20,21} which are key confounding factors that influence the accurate assessment of DNA methylation by obscuring the real peak at CpG sites. Recent studies have described several newly developed analysis methods^{20,22} on the basis of direct sequencing technology for methylation studies. Owing to their complicated procedures, their algorithms and workflow seem too difficult to be widely used.

To investigate whether it is feasible for direct PCR sequencing to accurately quantify DNA methylation with four-dye sequencing trace files by optimizing the corresponding procedures to eliminate the interfering factors, we first established a novel PCR program as described above. In addition, we used the ExoI/SAP enzymatic purification to effectively remove single bands, dNTP and unincorporated primers, which is easier, less expensive and requires smaller amounts of PCR products, compared with commercial PCR purification kits. Using these improved strategies, purified PCR products were sequenced with reverse primers (which are often better than forward primers), and the resulting signals in the sequencing chromatograms displayed evenly spaced peaks with uniform peak heights and unnoticeable baseline noise (Figures 2a, c and 3b, Supplementary Figure 3). We measured the peak height of C and T, and quantified the amount of methylation using the formula: % methylated C = $100\% \times \text{peak height C} / (\text{peak height C} + \text{peak height T})$.

Compared with the results of cloning-based sequencing in this study, the patterns of CpG methylation analyzed by the two methods were quite similar, with minor differences at some of the CpG sites (Figure 2). We selected another reliable method for quantification of DNA methylation, the pyrosequencing assay, to estimate the feasibility and precision of our method. Pyrosequencing is a sensitive and background-free assay, but it requires the CpG sites of interest very close to the sequencing primer, and the entire sequence length cannot be very long. Therefore, quantification of more CpG sites requires multiple primers and multiple reactions.^{18,27} We analyzed different methylation levels of the same PCR products with a length of 30 bp using our method and pyrosequencing. The two methods showed similar linearity and accuracy (Figure 4). Our strategy, however, generates minor background effects that are unavoidable in automated sequencing. However, given that the aim of quantitative evaluation of DNA methylation patterns was to determine whether DNA methylation has an essential role in the regulation of gene expression through comparison of DNA methylation levels among different samples, our method is sufficient to fulfill this purpose. In addition, the mean values of the coefficient of variation of our method seemed to be relatively high compared with those of the pyrosequencing

assay. One possible explanation for this is that the peak heights in trace files generated by direct PCR sequencing were affected by its neighbor base; therefore, the calculated methylation rates by our method were variable at different CpG sites with the same methylation level. Furthermore, systematic biases in the test system would lead to deviations from expected values and to a higher variance in the complete data, but it would still allow detection of relative differences in methylation rates at individual CpG positions.

In conclusion, we developed an attractive method for rapid analysis of DNA methylation through a series of optimization strategies and techniques aimed at solving the problems of direct bisulfite-PCR sequencing. The accuracy of our method is comparable with conventional screening methods such as pyrosequencing and cloning-based sequencing, while its speed and simplicity are superior to the latter two methods. Accordingly, the novel method is anticipated to be useful in screening large numbers of clinical samples across multiple genes.

Supplementary Information accompanies the paper on the Laboratory Investigation website (<http://www.laboratoryinvestigation.org>)

DISCLOSURE/CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was supported in part by the National Natural Science Foundation 30800622, Shanghai Science and Technology Research Program 09JC1402200, Shanghai leading Scientist for Public Health 08GWD07 and Shanghai Key Subject Project for Public Health 08GWZX0301. We thank J Gu (Shanghai Experimental Animal Center, China Academy of Sciences, Shanghai, China) for his help with animal studies.

1. Aoyama T, Okamoto T, Nagayama S, *et al*. Methylation in the core-promoter region of the chondromodulin-I gene determines the cell-specific expression by regulating the binding of transcriptional activator Sp3. *J Biol Chem* 2004;279:28789–28797.
2. Hashimshony T, Zhang J, Keshet I, *et al*. The role of DNA methylation in setting up chromatin structure during development. *Nat Genet* 2003;34:187–192.
3. Jones PA. DNA methylation and cancer. *Oncogene* 2002;21:5358–5360.
4. Leonhardt H, Cardoso MC. DNA methylation, nuclear structure, gene expression and cancer. *J Cell Biochem Suppl* 2000;79(Suppl 35):78–83.
5. Qiu GH, Salto-Tellez M, Ross JA, *et al*. The tumor suppressor gene DLEC1 is frequently silenced by DNA methylation in hepatocellular carcinoma and induces G1 arrest in cell cycle. *J Hepatol* 2008;48:433–441.
6. Tost J, Gut IG. Analysis of gene-specific DNA methylation patterns by pyrosequencing technology. *Methods Mol Biol* 2007;373:89–102.
7. Hashimoto K, Kokubun S, Itoi E, *et al*. Improved quantification of DNA methylation using methylation-sensitive restriction enzymes and real-time PCR. *Epigenetics* 2007;2:86–91.
8. Yang I, Park IY, Jang SM, *et al*. Rapid quantification of DNA methylation through dNMP analysis following bisulfite-PCR. *Nucleic Acids Res* 2006;34:e61.
9. Brena RM, Auer H, Kornacker K, *et al*. Accurate quantification of DNA methylation using combined bisulfite restriction analysis coupled with the Agilent 2100 Bioanalyzer platform. *Nucleic Acids Res* 2006;34:e17.
10. Kristensen LS, Mikeska T, Krypuy M, *et al*. Sensitive Melting Analysis after Real Time- Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. *Nucleic Acids Res* 2008;36:e42.
11. Eads CA, Danenberg KD, Kawakami K, *et al*. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res* 2000;28:E32.
12. Cottrell SE, Distler J, Goodman NS, *et al*. A real-time PCR assay for DNA-methylation using methylation-specific blockers. *Nucleic Acids Res* 2004;32:e10.
13. Thomassin H, Kress C, Grange T. MethylQuant: a sensitive method for quantifying methylation of specific cytosines within the genome. *Nucleic Acids Res* 2004;32:e168.
14. Choi JH, Park SH, Park J, *et al*. Site-specific methylation of CpG nucleotides in the hTERT promoter region can control the expression of hTERT during malignant progression of colorectal carcinoma. *Biochem Biophys Res Commun* 2007;361:615–620.
15. Hattermann K, Mehdorn HM, Mentlein R, *et al*. A methylation-specific and SYBR-green-based quantitative polymerase chain reaction technique for O6-methylguanine DNA methyltransferase promoter methylation analysis. *Anal Biochem* 2008;377:62–71.
16. Suzuki M, Yamada T, Kihara-Negishi F, *et al*. Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. *Oncogene* 2006;25:2477–2488.
17. Shimozaki K, Namihira M, Nakashima K, *et al*. Stage- and site-specific DNA demethylation during neural cell development from embryonic stem cells. *J Neurochem* 2005;93:432–439.
18. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc* 2007;2:2265–2275.
19. Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* 2001;29:E65–E65.
20. Paul CL, Clark SJ. Cytosine methylation: quantitation by automated genomic sequencing and GENESCAN analysis. *Biotechniques* 1996;21:126–133.
21. Myohanen S, Wahlfors J, Janne J. Automated fluorescent genomic sequencing as applied to the methylation analysis of the human ornithine decarboxylase gene. *DNA Seq* 1994;5:1–8.
22. Lewin J, Schmitt AO, Adorjan P, *et al*. Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplicates. *Bioinformatics* 2004;20:3005–3012.
23. Jiang MH, Fei J, Lan MS, *et al*. Hypermethylation of hepatic Gck promoter in ageing rats contributes to diabetogenic potential. *Diabetologia* 2008;51:1525–1533.
24. Bogdarina I, Murphy HC, Burns SP, *et al*. Investigation of the role of epigenetic modification of the rat glucokinase gene in fetal programming. *Life Sci* 2004;74:1407–1415.
25. Xu J, Christian B, Jump DB. Regulation of rat hepatic L-pyruvate kinase promoter composition and activity by glucose, n-3 polyunsaturated fatty acids, and peroxisome proliferator-activated receptor- α agonist. *J Biol Chem* 2006;281:18351–18362.
26. Qiu P, Soder GJ, Sanfiorenzo VJ, *et al*. Quantification of single nucleotide polymorphisms by automated DNA sequencing. *Biochem Biophys Res Commun* 2003;309:331–338.
27. Colella S, Shen L, Baggerly KA, *et al*. Sensitive and quantitative universal pyrosequencing methylation analysis of CpG sites. *Biotechniques* 2003;35:146–150.