Published online December 2, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 21 ***e168***
doi:10.1093/nar/gnh166

MethylQuant: a sensitive method for quantifying methylation of specific cytosines within the genome

Hélène Thomassin, Clémence Kress and Thierry Grange*

Institut Jacques Monod du CNRS, Universités Paris 6-7, Tour 43, 2 place Jussieu, 75251 Paris Cedex 05, France

Received September 28, 2004; Revised and Accepted November 5, 2004

ABSTRACT

Here we present MethylQuant, a novel method that allows accurate quantification of the methylation level of a specific cytosine within a complex genome. This method relies on the well-established treatment of genomic DNA with sodium bisulfite, which converts cytosine into uracil without modifying 5-methyl cytosine. The region of interest is then PCR-amplified and quantification of the methylation status of a specific cytosine is performed by methylation-specific real-time PCR with SYBR Green I using one of the primers whose 3′ end discriminates between the methylation states of this cytosine. The presence of a locked nucleic acid at the 3′ end of the discriminative primer provides the specificity necessary for accurate and sensitive quantification, even when one of the methylation states is present at a level as low as 1% of the overall population. We demonstrate that accurate quantification of the methylation status of specific cytosines can be achieved in biological samples. The method is high-throughput, cost-effective, relatively simple and does not require any specific equipment other than a real-time PCR instrument.

INTRODUCTION

Methylation of cytosine is an epigenetic DNA modification affecting gene expression. In vertebrates, it participates in the development, protection against intragenomic parasites, genomic imprinting, X chromosome inactivation and cancer (1–3). Cytosines within the dinucleotide CpG are methylated at the carbon-5 position after DNA synthesis by DNA methyltransferases (DNMT). The CpG dinucleotide is generally found in clusters called CpG islands. Methylation can participate in the control of gene expression in two ways. Dense methylation of CpG islands is associated with a repressive chromatin structure stabilized by methyl-binding proteins (MBDs). A more subtle effect is achieved when methylation of specific cytosines within binding sites of regulatory proteins interferes with their binding and action. The pattern of cytosine methylation is relatively stable since it is generally retained throughout cell division, but it can be dynamically reprogrammed in a regulated manner. The mechanisms involved in DNA demethylation and in the targeting of the enzymes modifying DNA methylation have not yet been identified clearly (4). Given the key roles of cytosine methylation, there has been a huge interest in the development of procedures for DNA methylation analyses (3). Sensitive, accurate, high-throughput and cost-effective methods should give a better definition of the role and control of this epigenetic modification, especially in pathological situations such as cancer.

Many different experimental approaches have been developed to allow either global, large-scale or specific analyses [for reviews see (3,5–7)]. The most popular approaches rely on bisulfite treatment of DNA (8). This treatment can be performed in such a way that, while cytosine is quantitatively deaminated to uracil, 5-methyl cytosine remains unmodified, thus allowing identification of the cytosine methylation status following PCR amplification. There are many possible variations in the subsequent DNA sequence detection methods used which achieve diverse levels of performance and adequacy to the analyses. Methylation-specific PCR (MSP) (9) is the most widely used assay for the detection of hypermethylation in CpG islands. It relies on the selective PCR amplification of sequences corresponding to either unmethylated or methylated DNA using primers that anneal specifically with either one of the DNA species. For methylation-specific annealing, each primer must contain sequences corresponding to at least two CpG dinucleotides. Thus, MSP can clearly identify the DNA molecules whose methylation statuses at multiple CpGs differ. This method allows sensitive detection of particular methylation patterns and appears very promising for the analysis of pathological samples. The MSP approach has been adapted to fluorescence-based real-time PCR analyses, thus providing both a quantitative and high-throughput technology. In the corresponding MethyLight and Heavy-Methyl approaches, a sequence-specific fluorescent detection probe is used, which hybridizes in a methylation-specific manner to a sequence corresponding to several additional CpGs (10,11). These approaches are thus best suited to the analysis of extensive changes in the methylation patterns within CpG islands, but they are not adapted to the detection of more subtle variations or to the analysis of the methylation of a single

*To whom correspondence should be addressed. Tel: +33 1 44275707; Fax: +33 1 44275716; Email: grange@ccr.jussieu.fr
Present address:
Hélène Thomassin, Institut Pasteur, CNRS URA 2578, 25-28 rue du Docteur Roux, 75724 Paris Cedex 15, France

specific cytosine either within or outside a CpG island. The quantitative analysis of the cytosine methylation levels at several independent proximal positions can be achieved using Pyrosequencing analysis of PCR products amplified in a methylation-independent manner (12,13). This procedure is not particularly sensitive, however, and does not perform well in the case of the detection of a population that represents <5–10% of the total, which limits its usefulness in the analysis of heterogeneous pathological samples and early events in epigenetic reprogramming.

Here, we present a novel fluorescence-based real-time PCR strategy, MethylQuant, which permits high-throughput quantification of the methylation status of a single specific cytosine in an accurate, sensitive and cost-effective manner.

MATERIALS AND METHODS

Sodium bisulfite conversion

To reduce the viscosity of genomic DNA and to facilitate the DNA denaturation required for efficient bisulfite conversion, DNA was digested with a restriction enzyme so that the restriction fragment encompassing the region of interest had a size of ~1 kb. Sodium bisulfite treatment was performed using the agarose bead method (14,15). Denatured DNA was embedded into low-melting-point agarose (SeaPlaque GTG, FMC); therefore, each agarose–DNA bead contained ~100 ng as follows: 900 ng of DNA in 30 µl of 0.2 M NaOH were incubated for 15 min at 50°C and mixed with 60 µl of 2% agarose solution at 50°C. Ten µl drops of this solution were pipeted in mineral oil at 4°C (not more than three drops per 200 µl of mineral oil per tube). After 10 min on ice, 200 µl of 5 M sodium bisulfite solution was added. This solution was prepared exactly as described previously (5). The beads were incubated in the dark at 50°C for 4 h. The bisulfite solution was then removed and the beads were washed four times for 15 min with 10 mM Tris–HCl, 1 mM EDTA, pH 8.0 (TE). The beads were then incubated twice in 500 µl of 0.2 M NaOH for 15 min at 20°C. Subsequently, 100 µl of 1 N HCl was added to the second incubation, and the beads were washed three times for 10 min with TE. Prior to PCR, the beads were washed twice in water for 15 min.

PCR amplification of bisulfite-treated DNA

Bisulfite-treated DNA was PCR amplified in the presence of tetramethylammonium chloride (TMAC) as described previously (5). One 10 µl bead was used in a 100 µl PCR reaction containing 67 mM Tris–HCl (pH 8.8), 16.6 mM $(NH_4)_2SO_4$, 0.5 mM TMAC, 6.7 mM $MgCl_2$, 500 µM dNTPs, 170 µg/ml DNase-free BSA, 400 nM of each primer and 2 U AmpliTaq DNA polymerase (Applied Biosystem). After an initial denaturation step for 4 min at 94°C, 32 cycles of amplification were performed as follows: 40 s at 94°C, 90 s at 55°C and 2 min at 72°C followed by a final extension step for 7 min at 72°C. PCR products were verified on an agarose gel before further processing. The primer sets that we used for the glucocorticoid-responsive unit (GRU) of the rat *Tat* gene were as follows:

lower strand set: CTTCTCAATATTCTCTATCACAA (−2615/−2593), ATTGAATAAAATTGTTTTATGGT (−2176/−2198);
upper strand set: TTTGTTGTATAGGATGTTTTAGT (−2510/−2488), ACCTACCTAACTTTAAAATTATC (−2115/−2137).

Quantitative real-time PCR

Real-time PCR analysis was performed using a LightCycler (Roche) and the LightCycler-FastStart DNA Master SYBR Green I mix (Roche). 3′-locked nucleic acid (LNA) primers were synthesized by Proligo. Reactions were performed in a final volume of 15 µl, adjusted to 4 mM $MgCl_2$ and containing 500 nM of each primer and 2 µl of DNA template. The template solutions were either a 10 000-fold dilution of the previously described PCR amplification of the samples or a serial 5-fold dilution of the matched (MAT) and mismatched (MIS) templates. These templates were prepared as follows: they were PCR amplified from individual plasmids resulting from the cloning in pGEM-T (Promega) of bisulfite-treated PCR-amplified DNA and selected according to their cytosine methylation pattern. The real-time PCR cycling conditions were as follows: 95°C for 8 min, followed by 50 cycles for 10 s at 95°C, 10 s at 56°C and 20 s at 68°C followed by fluorescence measurement. When standard discriminative (D) primers were used, the hybridization step was reduced to 5 s. Further, for comparative purposes, this shorter hybridization step was used to determine the properties of the primers presented in Table 1. The polymerization temperature was set to 68°C to allow accurate fluorescence measurements because of the low-melting temperature of the PCR products analyzed. Following PCR, a thermal melt profile was performed for amplicon identification. To determine the *Ct*,

Table 1. Comparison of the properties of the PCR amplifications with various D primers containing or not a 3′-LNA

	Regular primers			3′-LNA primers			Mismatch
	$\Delta Ct_{MIS-MAT}$	*E*	Bgd (%)	$\Delta Ct_{MIS-MAT}$	*E*	Bgd (%)	
U2$^-$ Unmet	7.2	0.79	1.5	12	0.78	0.10	A:C
U3$^-$ Unmet	8.8	0.75	0.7	13.3	0.76	0.05	A:C
L2$^+$ Unmet	6.8	0.72	2.5	14.4	0.66	0.07	A:C
L2$^+$ Met	N/A	N/A	N/A	11.7	0.78	0.12	G:T
L3$^-$ Unmet	4.5	0.76	7.9	14	0.72	0.05	T:G
L3$^-$ Met (56°C)	2.6	0.71	24.8	9.5	0.79	0.40	C:A
L3$^-$ Met (58°C)	4.3	0.65	11.6	10.3	0.76	0.30	C:A

The $\Delta Ct_{MIS-MAT}$ were determined for each D primer sets with identical amounts of MIS and MAT templates, the efficiency (*E*) of PCR amplification was determined using standard curves of MIS and MAT templates as described in Figure 2 and the Bgd value amplification with MIS template was calculated using Equation 1. The nature of the mismatch involving the 3′-most base of the D primer and the MIS template is indicated (primer:template). A primer annealing temperature of 56°C was used for all primer sets except for the L3$^-$ Met primer set that was analyzed at two temperatures as indicated.

the threshold level of fluorescence was set manually in the early phase of the PCR amplification. Each sample was analyzed at least in duplicate with either the non-discriminative (ND) or the discriminative (D) primer sets within the same run.

The primers that we used were named according to the following conventions: the cytosines included in the three CpG dinucleotides located within the rat tyrosine aminotransferase (*Tat*) GRU between −2419 and −2405 were numbered U1–U3 (−2419, −2413 and −2406, respectively), or L1–L3 (−2418, −2412 and −2405, respectively), depending on whether the strand was upper (U) or lower (L). (+) and (−) indicate the strands of the PCR product to which the primer was homologous. The D primers were named Unmet or Met to describe the methylation state of the cytosine to which they were matched. The opposing primer allowing PCR amplification when combined to either the ND or D primer was referred to as partner primer. The inosine substitutions are indicated by a small i.

U2$^-$ ND, ACACCCAAAAACCiACAAAC;
U2$^-$ Unmet, CACCCAAAAACCiACAAACA;
U3$^-$ Unmet, CACCACACCCAAAAACCA;
U2&3$^+$ partner, GGATGTTTTAGTTATTTTATTTGTA;
L2$^+$ ND, AAAACAAACAAATCCTACiTAATC;
L2$^+$ Unmet, AAACAAACAAATCCTACiTAATCA;
L2$^+$ Met, AAACAAACAAATCCTACiTAATCG;
L2$^-$ partner, TTTTTGAGATAAGGTTTTTTTAAG;
L3 ND, AGGGTTATATTATTATATTTAGAAAT;
L3$^-$ Unmet, GGGTTATATTATTATATTTAGAAATT;
L3$^-$ Met, GGGTTATATTATTATATTTAGAAATC;
L3$^+$ partner, AAAATTTACCAATCTCTACTATAC.

RESULTS

Principles of the MethylQuant technology

The quantification of the methylation level of a specific cytosine in bisulfite-treated DNA can be assimilated to the quantitative detection of a single nucleotide polymorphism (SNP) except that SNP detection does not usually require a method that is quantitative over a wide range. The basic principle of the MethylQuant procedure is that a methylation-specific PCR is performed with one of the primers harboring the methylation status-specific nucleotide at the 3′ end. Real-time PCR analysis allows quantification of the methylation status-specific product. In parallel, methylation-independent real-time PCR is performed to quantify the total product. The major problem that adversely affects this simple strategy is that a single mismatch at the 3′ end interferes with, but does not prevent, primer extension [(16) and references therein]. The product of this illegitimate primer extension event is indistinguishable from the product of the legitimate one and also participates in the PCR reaction. We have both quantified and minimized this illegitimate primer extension event to reduce it to ∼0.1% of the legitimate event, thus allowing quantification of methylation levels that are above this background value. The procedure involves two steps. First, the region containing the CpG of interest is amplified using primers that cover regions that are unlikely to contain methylated cytosines. Second, real-time PCR reactions are performed with either one of two primer sets that allow quantification of the total amplified material and of the fraction containing the nucleotide of interest.

The first amplification step prevents the introduction of biases that can result from the use of small amounts of starting material and from the heterogeneity of the quality of the various samples. Furthermore, it makes it possible to analyze sequences that are difficult to amplify and it offers more options in the design of the primers (see below). This amplification step is performed in a buffer containing TMAC which proved particularly useful for the amplification of difficult sequences, which are often encountered when using primers that do not cover regions containing methylated cytosine (5). Indeed, in this case, the primer content is very AT-rich owing to the conversion of cytosine into thymine, and TMAC improves the hybridization properties of such primers. The PCR product is then extensively diluted and real-time PCR analyses are performed using SYBR Green I detection of the PCR products. The use of SYBR Green I decreases the cost of this approach to a great extent when compared to the MethyLight procedures that use expensive sequence-specific oligonucleotide probes labeled with fluorophores. This cost saving becomes particularly significant when multiple positions are analyzed. For each sample, two real-time PCRs are performed using either one of two primer sets: (i) a ND primer set that allows quantification of the PCR product irrespective of the methylation status of the position of interest, and (ii) a D primer set that allows quantification of the PCR product corresponding to the specific methylation status of the position of interest (Figure 1B). The discriminative primer within the D set is designed in such a way that its most 3′ nucleotide is complementary to the position analyzed. Even though this is not essential, we obtained good results using, within the ND set, an almost identical primer shifted by one base pair so that its 3′ end directly abuts the position analyzed. The partner primer allowing PCR amplification is identical in the ND and D sets. This design helps to obtain ND and D primer sets that provide similar amplification efficiencies, which facilitates the quantification analysis. Furthermore, this ensures that the PCR products generated with each primer set fluoresce with similar efficiencies when bound by SYBR Green I. When designing the D primer, four different primers can be chosen to analyze methylation of a specific cytosine, two for the methylated state and two for the unmethylated state. Indeed, because the analysis is performed on PCR-amplified bisulfite-treated DNA, it is possible to design the D primer in such a way that it complements either of the two strands of the amplified product (Figure 1C). This confers a good flexibility to the primer design.

PCR-amplified bisulfite-treated DNA was cloned into a plasmid to generate homogeneous sequences corresponding to either the unmethylated or the methylated status of the cytosine of interest. PCR products generated from these plasmids were used to assess the ability of the D primer sets to discriminate effectively between the two templates. We refer to the template that is not perfectly hybridized to the D primer as the MIS template and to the other one as the MAT template. A scheme depicting a typical real-time PCR reaction is shown in Figure 1D. When the ND primer set is used, identical amounts of each template generally give rise to superimposable amplification curves. For the MethylQuant to be accurate, the amplification efficiencies of both the MAT and MIS templates with the ND primer set must be identical. When the D primer set is used to analyze the MAT template, it also gives

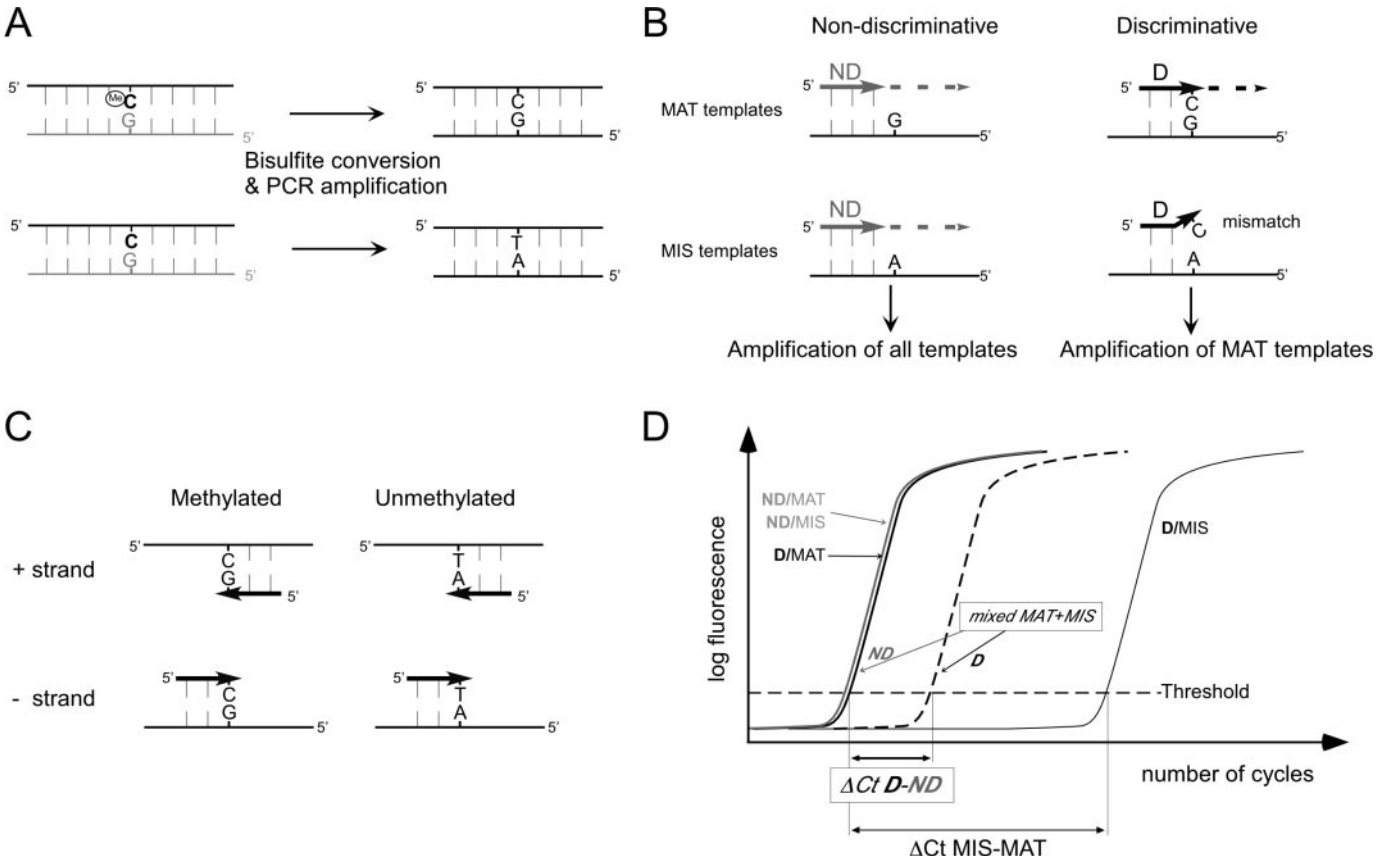

Figure 1. Principle of MethylQuant. (**A**) Nucleotide sequence modifications resulting from bisulfite conversion and PCR amplification. Only the cytosine of interest is represented. Since, following bisulfite conversion, the two DNA strands are no longer self-complementary, different primers are required. Thus, analysis of the complementary lower strand of the target sequence (gray lettering) is not shown. (**B**) Design of the D primers (black) or ND (gray) between the methylation states of the cytosine of interest. The MAT templates are matched to the D primer, whereas the 3′-most base of the D primer is mismatched to the MIS templates. **(C)** The four possible designs of a D primer. Following the first round of PCR amplification of the bisulfite-converted target sequence, either of the two strands (referred to as + or −) of the PCR product can be analyzed as represented. The D primer can be matched to the sequence corresponding to either the methylated or unmethylated cytosine. **(D)** Typical real-time PCR amplification profiles obtained with the ND (gray) and D (black) primer sets and various combinations of MIS and MAT templates. The log of the amount of PCR product is plotted as a function of the number of cycles. The dashed amplification curve corresponds to a sample containing a mixture of MIS and MAT templates. The threshold of PCR product synthesized allowing *Ct* determination is represented as a dashed line. The $\Delta Ct_{MIS-MAT}$ obtained by comparing the amplification of the MAT and MIS templates with the D primer set is represented as well as the ΔCt_{D-ND} obtained by comparing the amplification of a mixture of MIS and MAT templates with the ND and D primer sets.

rise to a superimposable amplification curve provided that the amplification efficiencies obtained with the ND and D primer sets are identical. This is not an absolute necessity, but it simplifies subsequent analyses (see below). Even when the 3′ end of the D primer is not hybridized with the template (MIS), illegitimate PCR amplification occurs but it is delayed with respect to amplification of the MAT template. This delay, measured in cycle numbers, is termed $\Delta Ct_{MIS-MAT}$, where *Ct*, the threshold cycle, is the fractional cycle number at which a threshold level of PCR product is detected. The value of $\Delta Ct_{MIS-MAT}$ determines the sensitivity of the approach, i.e. it will prevent the detection of a proportion of MAT template that is lower than the corresponding background (Bgd) value. The Bgd value (in %) can be calculated using the following equation:

$$\text{Bgd} = \frac{100}{(1+E)^{\Delta Ct_{MIS-MAT}}}, \quad \textbf{1}$$

where *E* is the efficiency of the PCR amplification varying between 0 and 1. For example, if the amplification efficiency is maximal, a $\Delta Ct_{MIS-MAT}$ of 5 cycles reveals that the Bgd of the procedure is $100/2^5$, i.e. ~3% of the initial mismatched population.

When the PCR product contains a mixture of MAT and MIS templates, with a percentage of MAT templates above the Bgd value, the PCR amplification curve will be found at an intermediate position (dashed curve in Figure 1D). The delay between the amplification obtained with the ND and D primer sets, ΔCt_{D-ND}, is related to the proportion of MAT templates within the population. If the efficiencies of the amplifications with the ND and D primer sets are rigorously identical, then the percentage of MAT templates can be calculated using the following equation:

$$\text{MAT} = \frac{100}{(1+E)^{\Delta Ct_{D-ND}}} - \text{Bgd}. \quad \textbf{2}$$

This equation can be applied as long as $\Delta Ct_{D-ND} \ll \Delta Ct_{MIS-MAT}$ in which case the Bgd value is negligible. When the values measured are close to that of the Bgd, then it is reasonable to refrain from quantifying.

The efficiencies of the PCR amplifications with the ND and D primer sets generally differ to some extent, in which case the percentage of MAT templates can be calculated using the following equation:

$$\mathrm{MAT} = 100\frac{(1+E_{\mathrm{ND}})^{Ct_{\mathrm{ND}}}}{(1+E_{\mathrm{D}})^{Ct_{\mathrm{D}}}} - \mathrm{Bgd}, \qquad \mathbf{3}$$

where Ct_{ND} and Ct_{D} are determined at an identical level of PCR product detected and E_{ND} and E_{D} are the efficiencies of the PCR amplifications with the ND and D primer sets, respectively.

Thus, quantification can be performed as long as the efficiency of the PCR amplification with the ND and D primer sets and the Bgd values are known. These values can be established using the standard procedures described in the real-time PCR manuals with titration curves of homogeneous MAT and MIS PCR products. Such PCR products are generated using cloned plasmids that are usually obtained during the preliminary analyses of the region of interest with the standard bisulfite, PCR, sequencing approach (8). The percentage of MAT template within a sample can also be determined from standard curves of MAT templates amplified with the ND and D primer sets using the analysis software provided with real-time PCR devices. In our experience, the greatest accuracy is achieved by determining the mean efficiency of each primer set in several independent experiments and by refraining from using the simplified Equation 2. Indeed, none of the ND and D primer sets we designed, amplified with a rigorously identical efficiency.

3′-Locked nucleic acid primers to improve methylation-specific PCR discrimination

The success of the MethylQuant procedure relies in its ability to discriminate effectively between the MAT and the MIS templates. The four types of 3′-primer:template mismatch that can be encountered in the MethylQuant procedure are G:T, A:C, C:A and T:G depending on the design of the D primer (see Figure 1C). The numerous studies performed to analyze the ability of allele-specific PCR to identify SNP reveal that these are among the least efficiently discriminated mismatches [(16,17) and references therein]. When developing MethylQuant, we analyzed cytosine methylation at four CpGs contained within the glucocorticoid-responsive unit (GRU) of the rat *Tat* gene (18). We tested a number of primers with various mismatches. Whatever the experimental conditions (variations in magnesium concentration, hybridization temperature and duration), the best conditions achieved only a moderate discrimination (Table 1). Depending on the primers, the Bgd amplification varies from 1 to 12% of the MIS template.

We obtained a distinct improvement in the discrimination ability when we used primers containing a single LNA at the 3′-terminal position. LNA is a nucleic acid analog with a 2′-O, 4′-C methylene bridge that locks the ribose moiety into a C3′-endo conformation, thus modifying its hybridization properties [(16) and references therein]. 3′-LNA primers allowed superior mismatch discrimination, bringing the Bgd amplification down to 0.4–0.05% of the MIS template depending on the primer considered (Table 1). A typical analysis is shown in Figure 2, which presents results obtained with standard curves of MAT and MIS templates analyzed with ND and D primer sets, the D primer either with or without a 3′-LNA (corresponding to the L3$^-$ Unmet primers in Table 1). Superimposed standard curves are obtained when both primer sets are complementary to the template (MAT template, Figure 2A). When the D primer set is not fully complementary (MIS template), the standard curve is shifted by the cycle number corresponding to its discriminative ability (ΔCt in Figure 2B). For the primer sets shown here, the mere presence of the 3′-LNA increased the $\Delta Ct_{\mathrm{MIS-MAT}}$, from 4 to 14 cycles, thereby decreasing the Bgd from 8 to 0.05% (L3$^-$ Unmet primers in Table 1). This improvement in the discrimination ability was observed for every primer set tested, even for those that were very unsatisfactory when regular primers were used (e.g. L3$^-$ Met in Table 1). The Bgd level with 3′-LNA primers was always inferior to 0.4% of the MIS template, and was generally <0.1%, thus rendering the MethylQuant procedure suitable for the quantification of low levels of methylation (or unmethylation) within a population.

When probing the methylation status of a specific cytosine with the 3′ end of a discriminative primer, it may happen that this primer hybridizes to a sequence containing other cytosines susceptible to methylation. To prevent the methylation of

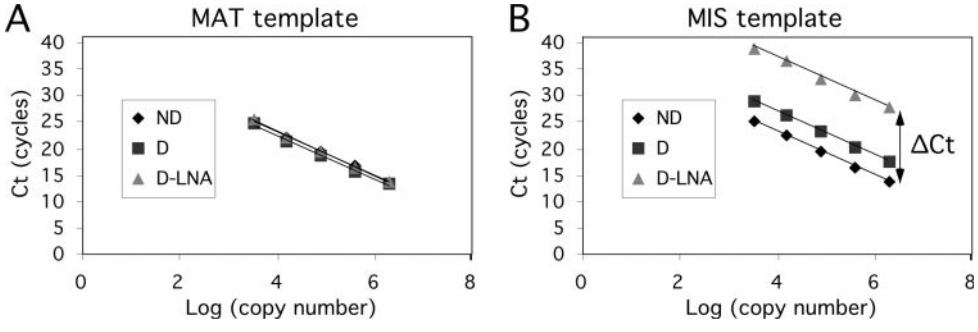

Figure 2. 3′-LNA primers show a superior discriminative ability. Real-time PCR analyses of serially diluted (5-fold increments) MAT (**A**) and MIS (**B**) templates with ND and D primer sets. The D primer contains or not a 3′-LNA as indicated. The results shown were obtained with the L3$^-$ primers described in Materials and Methods allowing assessment of cytosine −2405 of the rat *Tat* gene. The D primers (L3$^-$ Unmet) match to the − strand of the PCR product originating from the unmethylated lower strand of the *Tat* gene. The cycle number at threshold, *Ct*, is plotted as a function of the logarithm of the relative copy number of the templates. The slope of the linear regression analysis determined for each standard curve was used to calculate the amplification efficiency using the formula: $1+E=10^{-1/\mathrm{slope}}$. The ΔCt represented on (B) for the 3′-LNA D primer corresponds to the $\Delta Ct_{\mathrm{MIS-MAT}}$ as the standard curve obtained with amplification of the MIS template with the ND primer is superimposable to the standard curve obtained with amplification of the MAT template with the D-LNA primer [displayed in (A)].

these other cytosines from affecting the quantification of the position of interest, we used inosine at the corresponding positions within the ND and D primers, since it can base pair with any of the four conventional bases with approximately equal strength [(19), see sequence of the primers used in Materials and Methods]. The presence of inosine did not introduce a bias in the quantification since the MAT and MIS templates were amplified as efficiently with the inosine-containing ND primers. Thus, the combined use of inosine and 3′-LNA provided us with primers that made it possible to specifically assess the methylation status of any of the cytosines presently analyzed.

Accuracy and sensitivity of the MethylQuant procedure

To appreciate the capacity of the MethylQuant procedure to quantify cytosine methylation accurately within a heterogeneous sample, we performed dilution experiments by mixing different proportions of MAT and MIS templates. The means values of three independent measurements with standard deviations (SDs) were plotted against the actual ratios in the samples (Figure 3). In Figure 3A, the D primer allows quantification of a specific unmethylated cytosine, with mixtures of templates corresponding to a range of 0–100% of unmethylated DNA with intervals of 20%. Within this range, the method enabled us to quantify the input sample with a maximal SD of 15% of the measured value and the correct values were predicted with a maximal error of 15%. There was a linear relationship between the expected and measured allele frequencies with a slope of 0.98 and a correlation factor of 0.99. In Figure 3B, we used mixtures of templates corresponding to a range of 0–10% of unmethylated DNA to evaluate the ability of the procedure to measure low levels of demethylated cytosine, whereas in Figure 3C we used a D primer allowing quantification of the methylated state of the same cytosine at similar low levels. In both cases we obtained linear correlation curves with a correlation factor of 0.97. The low percentages were measured with a maximal SD of 15% of the measured value and the correct values were predicted with a maximal error of 20%. Thus, MethylQuant can be used to measure low levels of methylation or unmethylation in a DNA sample. When one of the two species is present at low levels, it is more accurate to measure directly the least abundant species than to deduce its levels from the measurements of the most abundant species.

Application of MethylQuant to the analysis of biological samples

We have used MethylQuant to analyze the demethylation of several cytosines triggered by the glucocorticoid receptor within the GRU of the rat *Tat* gene (18). We analyzed the methylation status of this regulatory region in DNA samples derived from rat hepatoma cells grown with glucocorticoids for various times. When using ligation-mediated PCR (LM-PCR) to analyze genomic DNA treated with hydrazine and piperidine, only unmethylated cytosines are detected (5). Glucocorticoid-triggered demethylation of three cytosines can be assessed on the gel shown in Figure 4A, where three extra bands appear over the course of a 4 day-glucocorticoid treatment. Even though it was possible to evaluate the extent of cytosine demethylation from the relative intensity of the bands, precise quantification was not possible owing to slight variations in the efficiencies of the various steps of the delicate multistep LM-PCR procedure and to the relatively high-level background. Thus, we used MethylQuant to measure the methylation levels of two cytosines using D primers that made it possible to detect both the methylated and unmethylated status of these cytosines (Figure 4B). MethylQuant confirmed the glucocorticoid-induced demethylation observed by the LM-PCR analysis and we are able to appreciate several properties that were not readily apparent: (i) ∼10% of the cytosines at each position were unmethylated in the cell population that had not been treated by glucocorticoids. This low level was not detected by LM-PCR due to the moderate Bgd at every position (Figure 4A). (ii) Glucocorticoid-induced DNA demethylation proceeds beyond 2 days of treatment, the methylation levels at the −2405 position decrease from ∼30% to ∼10% between 2 and 4 days of treatment, whereas those at the −2412 position decrease from ∼40% to ∼20%. These values were measured with the D primers allowing detection of the least abundant species, e.g. the unmethylated cytosines in the untreated cells, and the methylated cytosines in the cells treated for 2 and 4 days by glucocorticoid. This residual demethylation was not apparent with the LM-PCR

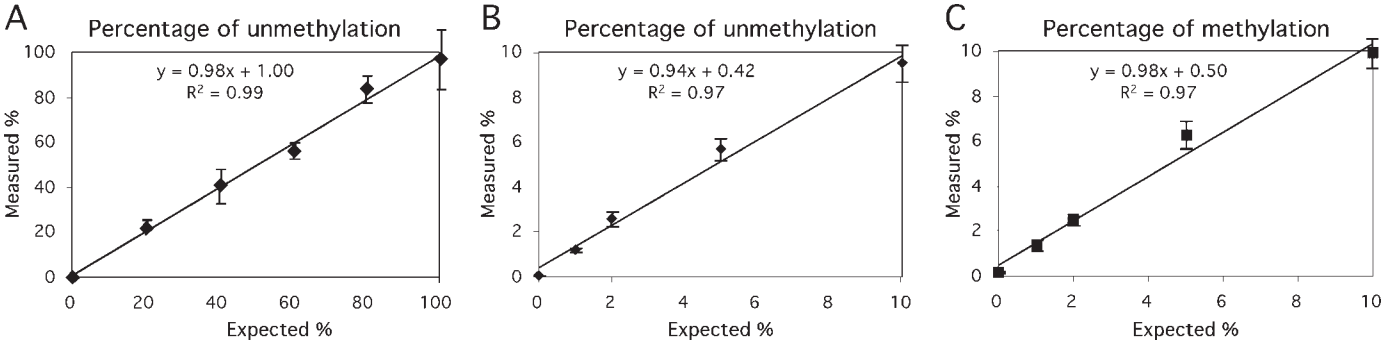

Figure 3. Accuracy and sensitivity of MethylQuant. PCR products obtained after amplification of bisulfite-converted lower strand of the methylated or unmethylated *Tat* GRU were mixed in different proportions. The amount of unmethylated or methylated cytosine −2405 was assessed using the 3′-LNA D primers L3^{-} Unmet (**A** and **B**) or L3^{-} Met (**C**). The measured percentage is plotted against the expected percentage as defined by the ratios of the two PCR products mixed together. For each data point at least three independent analysis were performed (SD are indicated by vertical bars). The equation of the linear regression curve as well as the correlation factor are indicated on each graph.

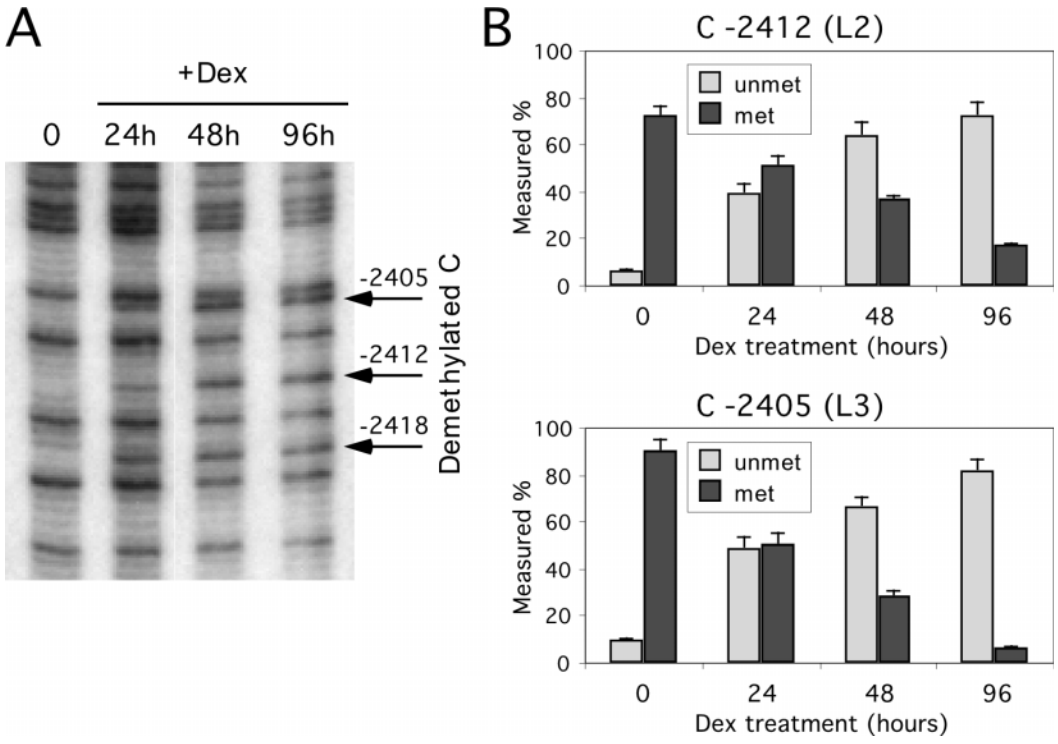

Figure 4. Demethylation of the *Tat* GRU induced by a prolonged glucocorticoid treatment. Rat hepatoma cells (H4II) were grown without (0) or with 10^{-7} M dexamethasone (+Dex) for 24, 48 and 96 h. The corresponding genomic DNA was purified and analyzed by either LM-PCR following hydrazine–piperidine treatment (5) (**A**) or by MethylQuant following bisulfite treatment (**B**). In (A), the cytosines that are subjected to demethylation are indicated on the right-hand side. In (B), the levels of either the unmethylated (gray bars) or the methylated (black bars) cytosines −2412 (upper graph) or −2405 (lower graph) were quantified using as D primers either the L2$^-$ Unmet, L2$^-$ Met, L3$^-$ Unmet or L3$^-$ Met primers, respectively.

analysis because the corresponding variations in the intensity of the demethylated cytosine bands were not consistent enough to be convincing. Similarly, with MethylQuant, the slight differences were not so convincing when the D primers were used to detect the most abundant species. As mentioned previously, it is more accurate to measure directly the least abundant species.

In an independent analysis of the demethylation of the *Tat* GRU, we used both MethylQuant and an extensive sequence analysis of cloned bisulfite-treated DNA. At the -2406 position in untreated cells, the level of unmethylated cytosine measured with MethylQuant was 10.0 ± 0.6% whereas the sequence analysis of 43 clones revealed a level of 9.3%. After a 10 h glucocorticoid treatment, the level measured with MethylQuant was 21.9 ± 0.6% whereas the sequence analysis of 77 clones revealed a level of 18.2%. Finally, after a 16 h glucocorticoid treatment, the level measured with MethylQuant was 38.6 ± 1.6% whereas the sequence analysis of 54 clones revealed a level of 38.9%. Thus, MethylQuant proved to be a sensitive and accurate method for measuring the levels of cytosine methylation at a given position in biological samples.

DISCUSSION

MethylQuant allows quantification of the methylation level of a specific cytosine using bisulfite-converted genomic DNA and real-time PCR. The quantification is based on the comparison of two PCRs performed with primer sets that amplify the target sequence either irrespective of methylation (ND primer set), or in a methylation-specific manner by means of a discriminative D primer whose 3′-most base hybridizes to the base corresponding to one methylation status. For quantification to be accurate (i) the D primer must correctly identify the methylation status; (ii) the efficiencies of amplification with the ND and D primer sets need to be determined and reproducible; and (iii) the amplification efficiencies of the templates corresponding to the different methylation levels must be the same. Unmodified D primers did not discriminate the different sequences well. We could obtain reasonable results by reducing the duration of the hybridization step to 5 s and by carefully optimizing the magnesium concentration and the hybridization temperature for each primer set. However, in these conditions, the reaction was sensitive to slight variations in the experimental conditions and sometimes differed depending on the batch of commercial reaction mixture that was used. The use of 3′-LNA primers considerably improved discrimination of the two sequences. It enabled us to lengthen slightly the hybridization step (10 s) which reduced the variability without any major effect on discrimination. We used ND and D primer sets that were as similar as possible, with a common partner primer and ND and D primers that differed only minimally (shifted by one base). Despite these precautions, the efficiencies of amplification were not absolutely identical in the vast majority of cases. To obtain a high level of precision in the measurements, it is necessary to determine accurately the amplification efficiency in several

independent experiments, to ensure that the experimental conditions do not cause significant fluctuations in the efficiency, and to use the mean efficiency for calculations. With 3′-LNA primers and the conditions described here, these requirements were not difficult to achieve. Just as in any real-time PCR analysis, with certain primers it is sometimes necessary to adjust temperature and magnesium concentration in order to fulfill these conditions. Finally, it is essential that the templates corresponding to targets methylated to different extents should be amplified with identical efficiencies. Since each unmethylated CpG dinucleotide gives rise to templates containing 2 A:T base pairs instead of 2 G:C base pairs, the melting temperatures of the templates differ as a function of the methylation levels of the targets. This could be responsible for a PCR bias (20). To check that the amplification efficiency is not affected by this phenomenon, we analyzed amplification with ND primer sets of templates corresponding to fully methylated or demethylated targets. With 106–145 bp long PCR fragments containing either two or three CpG, we measured with SYBR Green I, differences in the melting temperature of ~1°C depending on the methylation levels, but we did not observe any differences in the amplification efficiencies. This could be due to the fact that the melting temperatures of these templates are much lower than that of the denaturation step (below 75°C versus 94°C). Similarly, we verified that such amplification biases did not occur during the first PCR amplification step performed in the TMAC-containing buffer. Nevertheless, we recommend verification of the absence of any such amplification biases when novel sequences are analyzed, in particular if they contain a higher density of CpG. Provided that these specific requirements has been fulfilled, MethylQuant proved to be an accurate method for measuring cytosine methylation levels. We obtained almost identical results using MethylQuant and sequence analysis of ~50–75 clones. When high precision in the determination of the methylation levels of a limited number of cytosines is required, MethylQuant is much more time- and cost-effective than the sequence analysis of multiple clones.

Since, upon bisulfite treatment and PCR amplification, the determination of the methylation status of cytosines is akin to the analysis of nucleotide sequence differences, a number of DNA sequence analysis procedures have been adapted to the study of DNA methylation. Several quantification procedures have been described. With methylation-sensitive single nucleotide primer extension (Ms-SNuPE), the query primer terminates immediately 5′ of the nucleotide to be assayed and a single nucleotide is incorporated (21). Ms-SNuPE employs radiolabeled dNTPs and requires the purification of the PCR products before the primer extension reaction can be carried out and it is therefore more labor intensive and requires a greater number of analysis steps than MethylQuant. Furthermore, it is not clear that it is as sensitive as MethylQuant. Several techniques already developed to explore SNPs, in particular SNaP-shot and Pyrosequencing, have also been adapted to cytosine methylation analysis (12). SNaP-shot is based on a single nucleotide primer extension approach with fluorescently labeled ddNTPs while Pyrosequencing is based on a real-time sequencing technique. Pyrosequencing was shown to be more reliable and accurate than SNaP-shot (12) and has been used by several independent laboratories (13). All the same, both SnaP-shot and Pyrosequencing require a specific and expensive laboratory setup. Furthermore, they are not sensitive enough to measure low levels of a methylation state. In contrast to MethylQuant, Pyrosequencing allows simultaneous analyses of neighboring cytosines and might be more appropriate for measuring relatively high levels of cytosine methylation states at several different positions. The MethyLight and HeavyMethyl procedures are similar to MethylQuant since they also use fluorescence-based real-time PCR (10,11), but these procedures are more expensive than MethylQuant because they depend on real-time detection probes with a 5′-fluorescent reporter dye and a 3′ quencher dye instead of the more economical SYBR Green I and 3′-LNA modification. Furthermore, they require multiple cytosines to differ in their methylation state. When this requirement is fulfilled, they might be more sensitive than MethylQuant, in particular for the detection of low levels of methylated CpG islands within biological samples. This can be useful for the early detection in biopsies of tumor cells harboring the global methylation of a given CpG island. However, it remains to be determined whether early tumor cells already harbor such global methylation of all the CpGs within an island. Indeed, methylation of specific cytosines within the binding site of a key transcriptional activator might be sufficient to downregulate a tumor-suppressor gene and to contribute thus to early steps of tumorigenesis. Since MethylQuant is particularly suitable for the quantification of the methylation levels of single cytosines, it should prove useful for the analysis of the involvement of the methylation of key residues in tumorigenesis. We also found that its exceptional sensitivity and accuracy were a great advantage when we analyzed the mechanism of modification of epigenetic memory in response to transcriptional activation (C. Kress, H. Thomassin and T. Grange, manuscript in preparation).

ACKNOWLEDGEMENTS

We thank C. Brossas and A. Kropfinger for critical reading of the manuscript. This work was supported in part by the CNRS and by grants from the Association de Recherche sur le Cancer. C.K. was supported by fellowships from the CNRS and the Ligue Nationale contre le Cancer.

REFERENCES

1. Jones,P.A. and Baylin,S.B. (2002) The fundamental role of epigenetic events in cancer. *Nature Rev. Genet.*, **3**, 415–428.
2. Jaenisch,R. and Bird,A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genet.*, **33** (Suppl.), 245–254.
3. Laird,P.W. (2003) The power and the promise of DNA methylation markers. *Nature Rev. Cancer*, **3**, 253–266.
4. Kress,C., Thomassin,H. and Grange,T. (2001) Local DNA demethylation: how could it be achieved precisely, quickly and safely? *FEBS Lett.*, **494**, 135–140.
5. Thomassin,H., Oakeley,E.J. and Grange,T. (1999) Identification of 5-methylcytosine in complex genomes. *Methods*, **19**, 465–475.
6. Fraga,M.F. and Esteller,M. (2002) DNA methylation: a profile of methods and applications. *Biotechniques*, **33**, 632, 634, 636–649.
7. Liu,Z.J. and Maekawa,M. (2003) Polymerase chain reaction-based methods of DNA methylation analysis. *Anal. Biochem.*, **317**, 259–265.
8. Frommer,M., McDonald,L.E., Millar,D.S., Collis,C.M., Watt,F., Grigg,G.W., Molloy,P.L. and Paul,C.L. (1992) A genomic sequencing

protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl Acad. Sci. USA*, **89**, 1827–1831.
9. Herman,J.G., Graff,J.R., Myohanen,S., Nelkin,B.D. and Baylin,S.B. (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl Acad. Sci. USA*, **93**, 9821–9826.
10. Eads,C.A., Danenberg,K.D., Kawakami,K., Saltz,L.B., Blake,C., Shibata,D., Danenberg,P.V. and Laird,P.W. (2000) MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res.*, **28**, E32.
11. Cottrell,S.E., Distler,J., Goodman,N.S., Mooney,S.H., Kluth,A., Olek,A., Schwope,I., Tetzner,R., Ziebarth,H. and Berlin,K. (2004) A real-time PCR assay for DNA-methylation using methylation-specific blockers. *Nucleic Acids Res.*, **32**, E10.
12. Uhlmann,K., Brinckmann,A., Toliat,M.R., Ritter,H. and Nurnberg,P. (2002) Evaluation of a potential epigenetic biomarker by quantitative methyl-single nucleotide polymorphism analysis. *Electrophoresis*, **23**, 4072–4079.
13. Tost,J., Dunker,J. and Gut,I.G. (2003) Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing. *Biotechniques*, **35**, 152–156.
14. Olek,A., Oswald,J. and Walter,J. (1996) A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res.*, **24**, 5064–5066.
15. Hajkova,P., el-Maarri,O., Engemann,S., Oswald,J., Olek,A. and Walter,J. (2002) DNA-methylation analysis by the bisulfite-assisted genomic sequencing method. *Methods Mol. Biol.*, **200**, 143–154.
16. Latorra,D., Campbell,K., Wolter,A. and Hurley,J.M. (2003) Enhanced allele-specific PCR discrimination in SNP genotyping using 3′ locked nucleic acid (LNA) primers. *Hum. Mutat.*, **22**, 79–85.
17. Ayyadevara,S., Thaden,J.J. and Shmookler Reis,R.J. (2000) Discrimination of primer 3′-nucleotide mismatch by *Taq* DNA polymerase during polymerase chain reaction. *Anal. Biochem.*, **284**, 11–18.
18. Thomassin,H., Flavin,M., Espinás,M.L. and Grange,T. (2001) Glucocorticoid-induced DNA demethylation and gene memory during development. *EMBO J.*, **20**, 1974–1983.
19. Martin,F.H., Castro,M.M., Aboul-ela,F. and Tinoco,I.,Jr (1985) Base pairing involving deoxyinosine: implications for probe design. *Nucleic Acids Res.*, **13**, 8927–8938.
20. Warnecke,P.M., Stirzaker,C., Melki,J.R., Millar,D.S., Paul,C.L. and Clark,S.J. (1997) Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res.*, **25**, 4422–4426.
21. Gonzalgo,M.L. and Jones,P.A. (1997) Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.*, **25**, 2529–2531.