Array-based analysis of genomic DNA methylation patterns of the tumour suppressor gene *p16^INK4A* promoter in colon carcinoma cell lines

Cora Mund, Verena Beier[1,*], Peter Bewerunge[2], Michael Dahms[3], Frank Lyko and Jörg D. Hoheisel[1]

Division of Epigenetics and [1]Division of Functional Genome Analysis and [2]Division of Intelligent Bioinformatics Systems, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany and [3]febit AG, Käfertaler Strasse 190, 68167 Mannheim, Germany

ABSTRACT

Aberrant DNA methylation at CpG dinucleotides can result in epigenetic silencing of tumour suppressor genes and represents one of the earliest events in tumourigenesis. To date, however, high-throughput tools that are capable of surveying the methylation status of multiple gene promoters have been restricted to a limited number of cytosines. Here, we present an oligonucleotide microarray that permits the parallel analysis of the methylation status of individual cytosines, thus combining high throughput and high resolution. The approach was used to study the CpG island in the promoter region of the tumour suppressor gene *p16^INK4A*. In total, 876 oligonucleotide probes of 21 nt in length were used to inspect the methylation status of 53 CpG dinucleotides, producing correct signals in colorectal cancer cell lines as well as control samples with a defined methylation status. The information was validated by established alternative methods. The overall methylation pattern was consistent for each cell line, while different between them. At the level of individual cytosines, however, significant variations between individual cells of the same type were found, but also consistencies across the panel of cancer cell lines were observed.

INTRODUCTION

In the human genome, ~4% of the cytosine residues are modified by methylation at their carbon-5 position. DNA methylation plays a key role in genomic imprinting, the maintenance of genome integrity and has been implicated as an important factor in the development of cancer (1–3). Alterations in the epigenetic pattern are among the earliest and most common events in tumourigenesis (4,5). Almost all methylation takes place at CpG dinucleotide: a cytosine that is located 5′ to a guanine (6). While this dinucleotide is generally underrepresented in the human genome sequence, there are local regions—CpG islands—which are comparatively rich in unmethylated CpG sites. More than 60% of all known human genes contain CpG islands. Extensive methylation of these sites frequently leads to gene silencing (7). Aberrant hypermethylation of CpG islands in cancer cells is responsible for the transcriptional inactivation of many genes, including tumour suppressor genes (4).

Several methods exist to detect changes in the DNA methylation pattern. Most of these techniques use bisulfite treatment to uncover the methylation status. Sodium bisulfite induces methylation-dependent single nucleotide polymorphisms by converting unmethylated cytosine to uracil and, upon PCR amplification, to thymine. 5-methylcytosine is highly resistant to sodium bisulfite treatment and thus becomes amplified as cytosine. The most common assay to detect DNA methylation changes is methylation-specific PCR (MSP) (8). Two primer sets—one specific for the C-containing sequence, the other binding to T-containing DNA—are used to discriminate between the methylated and unmethylated status of genomic DNA. Also, the combined bisulfite restriction analysis (COBRA) (9), uses sodium bisulfite treatment, which is followed by PCR-amplification of the relevant regions, restriction digestion and quantification of the resulting restriction fragments. Variation in methylation is indicated by the creation of new or the retention of pre-existing restriction sites. Both methods examine only a very limited number of cytosines and a single gene per assay. In contrast, sequencing of

---

bisulfite-deaminated and PCR-amplified DNA (10) produces results for every CpG in the region of interest. However, read length is limited and the process is both cost- and time-consuming.

As an alternative, oligonucleotide microarrays represent a powerful tool for inexpensively generating a large amount of information in a single experiment, while nevertheless working at a single-base resolution. For the same reason, microarrays are used in other genotyping assays (11). Recently, microarrays were applied to study methylation in leukaemias (12), Non-Hodgkin's lymphoma (13) and gastric tumour (14). The results provided evidence that defined DNA methylation profiles might represent an important characteristic for tumour classification (12). All these studies, however, were limited by their restriction to a small number of CpG dinucleotides in a few selected genes. Another array-based approach indicated the potential for a genome-wide screening of CpG islands in promoter regions (15). With this array, the methylation scores of 276 anonymous CpG islands were analysed in a group of breast cancer cell lines. However, the methylation status of defined CpG dinucleotides could not be determined.

To date, little was done to establish high throughput methods that allow a detailed analysis of the DNA methylation pattern of individual CpG islands. Here, we present the design, establishment and validation of a microarray system that is capable of surveying the methylation status of multiple CpGs individually. Oligonucleotide microarrays were produced by light-directed, mask-free *in situ* synthesis using the *Geniom One* technology of febit AG (16). This array format allows the flexible production of oligonucleotide arrays with probes for every CpG within genomic regions of interest. We focussed our analysis on the $p16^{INK4A}$/ *CDKN2A/INK4a* tumour suppressor gene, which is located in the chromosomal region 9p21 and encodes the cyclin-dependent kinase inhibitor 2A. This gene is either deleted or mutated in a wide range of cancers and was found to be silenced by CpG methylation in many tumour types (17). After hybridization of DNA samples from three colorectal cell lines and analysing ~29 000 signals produced in each experiment, the methylation status was deduced from the chip data and independently validated with conventional techniques, documenting the detailed resolution of the assay. From these studies, we could also define oligonucleotide probes that exhibited an especially good ability to discriminate between methylated and unmethylated DNA. Selection criteria could be deduced from this information, which will translate into improved oligomer design for expanded DNA methylation studies in cancer research.

MATERIALS AND METHODS

Cell lines

The colon carcinoma cell lines HCT116, SW480 and SW48 were cultured under standard conditions. HCT116 cells were grown in McCoy's 5a medium (Invitrogen, Karlsruhe, Germany) supplemented with 10% fetal calf serum (FCS), while RPMI1640 medium (Invitrogen) supplemented with 10% FCS and 5% L-glutamine was used for SW480 and SW48 cells.

DNA preparation

Genomic DNA was prepared from cells with the DNeasy tissue kit of Qiagen (Hilden, Germany) according to the manufacturer's protocol and digested with EcoRI. An unmethylated control target was produced by PCR-amplifying genomic DNA with forward primer d(AAACACGCCTTTGCTGG-CAGG) and reverse primer d(CCACGCGCGTACAGATC-TCT). A fully methylated target was generated by *in vitro* methylation of 700 ng of the resulting PCR fragment with 8 U of SssI methylase (New England Biolabs, Frankfurt, Germany) at 37°C for 2 h. Partially methylated targets were generated by *in vitro* methylation with 8 U of bacterial methylase M.HpaII or M.HhaI (New England Biolabs, Frankfurt, Germany), respectively, at 37°C for 2 h. To determine the methylation efficacy, each template was digested with the corresponding restriction enzyme, HpaII or HhaI (New England Biolabs, Frankfurt, Germany), for 1 h at 37°C and analysed by agarose gel electrophoresis.

Bisulfite treatment

DNA was treated with sodium bisulfite as described elsewhere (10). In brief, 700 ng DNA was denatured in 50 µl of 0.2 mM NaOH at 37°C for 20 min; 30 µl of freshly prepared 10 mM hydroquinone (Sigma, Heidelberg, Germany) and 520 µl of 3 M sodium bisulfite (Sigma) at pH 5 were added. Incubation was under a layer of mineral oil at 55°C for 14 h. The DNA was purified using the Gene Clean kit (Qbiogene, Heidelberg, Germany) and eluted with 100 µl water. Finally, the DNA was desulfonated by a 20 min treatment in 0.3 M NaOH at 37°C, followed by an ethanol precipitation. The DNA pellet was resuspended in 1 mM Tris–HCl, pH 8, and stored at −20°C.

Methylation-specific PCR (MSP) and combined bisulfite restriction analysis (COBRA)

MSP was performed as described previously (8). About 100 ng of bisulfite-modified DNA was amplified with primer pairs specific to unmethylated [U primer pair: d(TTATTAGAGG-GTGGGGTGGATTGT), d(CAACCCCAAACCACAACCA-TAA)] or methylated DNA sequences [M primer pair: d(TTATTAGAGGGTGGGGCGGATCGC), d(GACCCCGA-ACCGCGACCGTAA)], respectively. After an initial denaturation at 95°C for 3 min, DNA was amplified in 30 cycles of 95°C for 30 s, 60°C (U primer pair) or 65°C (M primer pair) for 30 s and 72°C for 45 s, followed by a final extension at 72°C for 5 min. COBRA was performed according to a previously published protocol (9). In brief, sodium bisulfite-treated DNA was PCR-amplified, digested with the restriction enzyme BstUI (New England Biolabs) and separated on a 1.5% agarose gel.

Array design

In the analysis of 53 CpG dinucleotides of the p16 promoter region, 252 control and 876 analysis probes were used. The 252 oligonucleotide probes acted as positive controls for the occurrence of synthesis and hybridization or represented entirely independent sequences, which indicated the background signal of unspecific binding. The analysis probe set reflected the methylation status of every CpG dinucleotide. If an oligonucleotide probe covered more than one CpG site, probes for every possible combination of methylated and
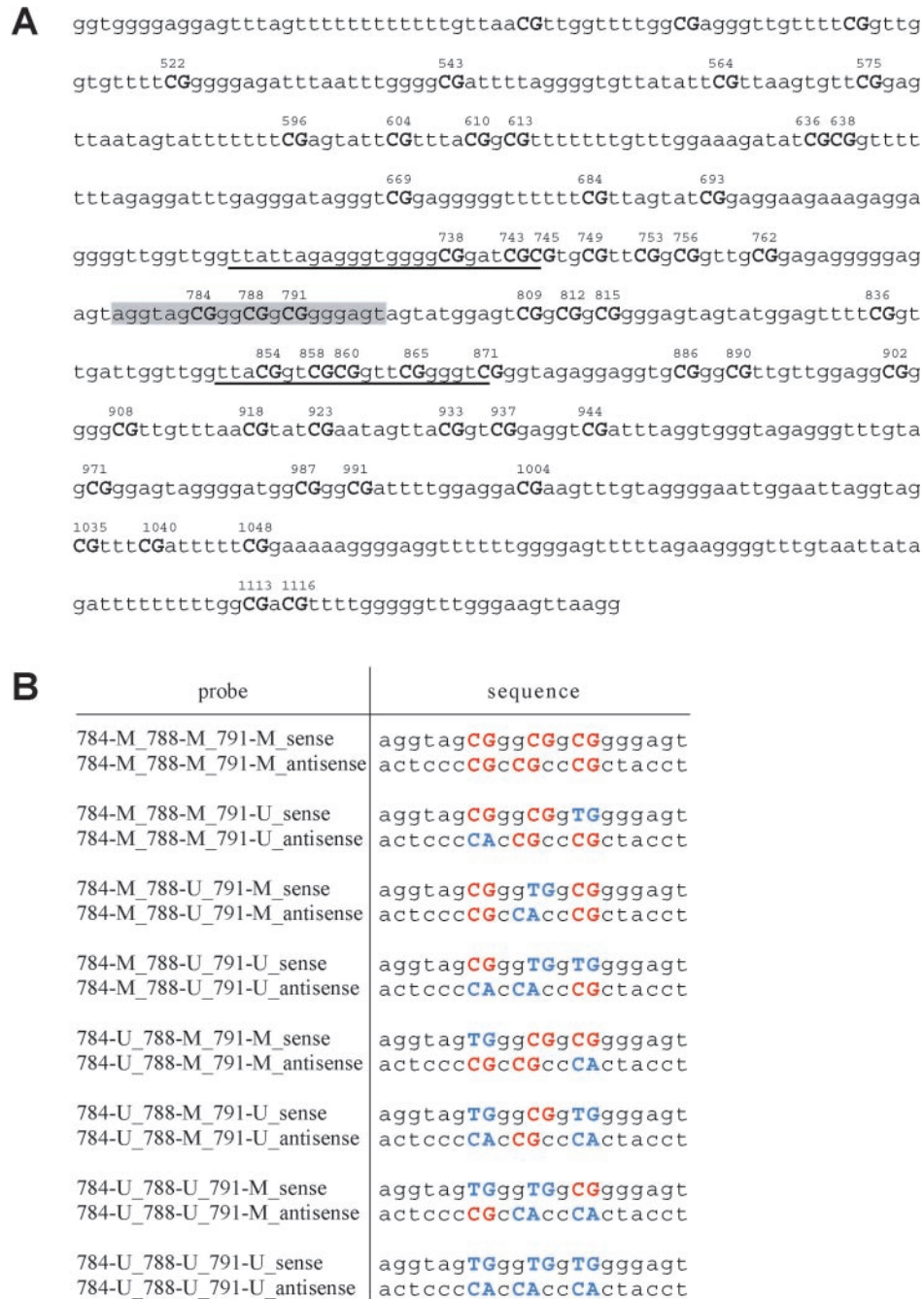
A
```
ggtggggaggagtttagtttttttttttttttgttaaCGttggtttttggCGagggttgtttttCGgttg

      522                     543                          564         575
gtgttttCGggggagatttaatttggggCGattttaggggtgttatattCGttaagtgttCGgag

      596      604    610 613
ttaatagtatttttttCGagtattCGtttaCGgCGttttttttgtttggaaagatatCGCGgtttt

               669              684        693
tttagaggatttgagggatagggtCGgaggggggtttttttCGttagtatCGgaggaagaaagagga

                    738    743 745 749   753 756    762
ggggttggttggttattagagggtggggCGgatCGCGtgCGttCGgCGgttgCGgagaggggggag

        784  788 791                    809 812 815                  836
agtaggtagCGggCGgCGgggagtagtatggagtCGgCGgCGgggagtagtatggagttttCGgt

          854 858 860  865    871                886   890          902
tgattggttggttaCGgtCGCGgttCGgggtCGggtagaggaggtgCGggCGttgttggaggCGg

     908        918    923        933  937    944
gggCGttgtttaaCGtatCGaatagttaCGgtCGgaggtCGatttaggtgggtagagggtttgta

    971               987  991           1004
gCGggagtaggggatggCGggCGattttggaggaCGaagtttgtaggggaattggaattaggtag

  1035   1040      1048
CGtttCGattttttCGgaaaaaggggaggtttttttggggagttttttagaaggggtttgtaattata

          1113 1116
gattttttttttggCGaCGtttttgggggtttgggaagttaagg
```

B

probe	sequence
784-M_788-M_791-M_sense	aggtag**CG**gg**CG**g**CG**gggagt
784-M_788-M_791-M_antisense	actccc**CG**c**CG**cc**CG**ctacct
784-M_788-M_791-U_sense	aggtag**CG**gg**CG**g**TG**gggagt
784-M_788-M_791-U_antisense	actccc**CA**c**CG**cc**CG**ctacct
784-M_788-U_791-M_sense	aggtag**CG**gg**TG**g**CG**gggagt
784-M_788-U_791-M_antisense	actccc**CG**c**CA**cc**CG**ctacct
784-M_788-U_791-U_sense	aggtag**CG**gg**TG**g**TG**gggagt
784-M_788-U_791-U_antisense	actccc**CA**c**CA**cc**CG**ctacct
784-U_788-M_791-M_sense	aggtag**TG**gg**CG**g**CG**gggagt
784-U_788-M_791-M_antisense	actccc**CG**c**CG**cc**CA**ctacct
784-U_788-M_791-U_sense	aggtag**TG**gg**CG**g**TG**gggagt
784-U_788-M_791-U_antisense	actccc**CA**c**CG**cc**CA**ctacct
784-U_788-U_791-M_sense	aggtag**TG**gg**TG**g**CG**gggagt
784-U_788-U_791-M_antisense	actccc**CG**c**CA**cc**CA**ctacct
784-U_788-U_791-U_sense	aggtag**TG**gg**TG**g**TG**gggagt
784-U_788-U_791-U_antisense	actccc**CA**c**CA**cc**CA**ctacct

Figure 1. Structure of the analysed CpG island in the *p16* promoter region. (**A**) The DNA sequence of the CpG island is shown. The CpG dinucleotides are numbered according to their position in the sequence. Because a longer sequence was used for the initial array design, numbering does not start at position 1. The binding positions of the primers used in the MSP analysis are underlined. As an example for a sequence area covered by array-bound oligonucleotide probes, a 21mer sequence is marked in grey, which includes the CpG dinucleotides 784, 788 and 791. (**B**) Design of oligonucleotide probes. The 16 probe sequences are shown, which were required to assay all possible combinations of methylated (M) and unmethylated (U) CpG dinucleotides in this 21mer sequence.

unmethylated cytosines were generated (e.g. Figure 1). For n CpG dinucleotides within a region covered by an oligomer probe, there are 2^n different probe sequences. Additionally, for every sense probe, the reverse complementary antisense probe was generated. For an evaluation of reproducibility, the analysis probe set was synthesized eight times on each array. The sequences are available at www.dkfz.de/funct_genome/epi-p16-oligos (and also as Supplementary Data).

Array production

Oligonucleotide arrays were generated by photo-controlled *in situ* synthesis as previously described (16) using the *Geniom One* system and DNA processors provided by febit AG (now available from febit biotech, Heidelberg, Germany). Each DNA processor consists of four individually accessible micro-channels, each of which is referred to as array in this manuscript. Since in our analysis every array contained

7260 probes, the whole DNA processor consisted of 29 040 oligonucleotides. They were synthesized in parallel in all four arrays using standard DNA synthesis reagents (Proligo, Hamburg, Germany) and 3′-phosporamidites carrying a 5′-photolabile group (18,19).

Labelling, hybridization and signal detection

For each hybridization, 50–100 ng PCR product was labelled by random hexamer priming (20) in the presence of biotin-dUTP (Roche, Mannheim, Germany). Prior to hybridization, a prehybridization was performed in 100 mM 2-[*N*-morpholino]ethanesulfonic acid (MES), 0.9 M NaCl, 20 mM Na_2EDTA, 0.01% (v/v) Tween-20, 1% BSA (Sigma) at room temperature for 15 min. To each array, 5–10 ng/µl of biotin-labelled DNA were hybridized in the same buffer but with only 0.5 mg/ml BSA and supplemented with 0.1 mg/ml herring sperm DNA (Promega, Mannheim, Germany). Hybridization was at 45°C for 3 h, followed by a first washing step at 25°C with 6× SSPE (0.9 M NaCl, 60 mM NaH_2PO4, pH 7.4 and 6 mM Na_2EDTA) and a second washing step at 45°C with 0.5× SSPE supplemented with 0.005% (v/v) Triton X-100. Fluorescence-staining was performed with 0.1 mg/ml streptavidin-R-phycoerythrin-conjugate (Molecular Probes, Leiden, The Netherlands) in 6× SSPE at 25°C for 15 min. Subsequently, the arrays were rinsed with 6× SSPE at 25°C. Signal detection was done using the internal CCD-camera system of the *Geniom One* instrument. For detection of the phycoerythrine chromophore, the Cy3 filter was employed. The duration of signal integration was determined automatically by the instrument's software. Typical integration times were 350–500 ms.

Bisulfite sequencing

After bisulfite treatment as described above, PCR amplification was carried out in 20 µl Thermoprime polymerase and ReddyMix reaction buffer (Abgene, Wiesenbach, Germany) in the presence of 1.25 mM of each dNTP and 400 nM of the primers d(CGTGGGGAGGAGTTTAGTTTTTTTT) and d(CCTTAACTTCCCAAACCCCCAAAA). An initial incubation at 95°C for 3 min was followed by 30 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 45 s and a final extension at 72°C for 5 min. The amplification product was purified from a 1.5% agarose gel and cloned with the TA-TOPO Cloning kit (Invitrogen). Sequencing on the plasmid DNA was performed by GATC (Constance, Germany).

RESULTS

p16 methylation analysis by MSP and COBRA

For DNA methylation analysis, we used three colorectal cancer cell lines with different *p16* methylation patterns. HCT116 cells have been shown to contain both an unmethylated and a hypermethylated allele (21). In the SW 480 cell line, the promoter of the *p16* gene has been reported to be methylated (22), while the *p16* methylation status of SW48 cells has not been published so far. In order to establish reliable conditions for bisulfite deamination, we analysed *p16* methylation by using methylation-specific PCR (MSP) and combined bisulfite restriction analysis (COBRA). Genomic DNA was treated
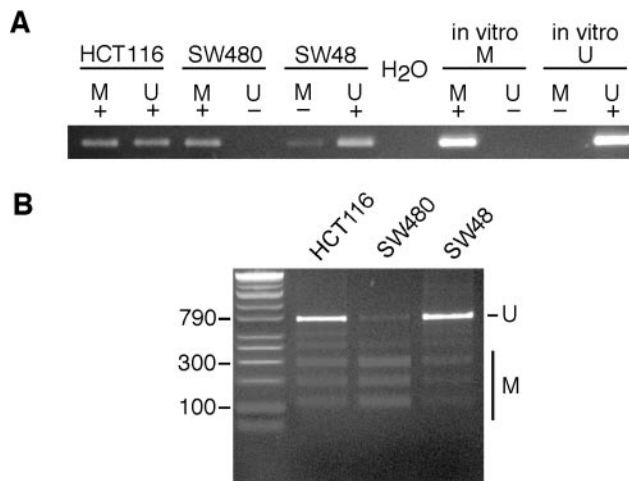

Figure 2. *p16* methylation analysis of genomic DNA from HCT116, SW480 and SW48 cells. (**A**) Methylation analysis by MSP. Two primer pairs were applied that distinguish between methylated (M) and unmethylated (U) DNA. In HCT116 cells, amplification occurred in both reactions, producing bands of equal intensity. This is indicative for the existence of methylated (M) and unmethylated (U) DNA in similar quantities. For SW480 cells, the MSP resulted in an M band only. Analysis of material made from SW48 cells exhibited a prominent U band and a much weaker M band. *In vitro* methylated and unmethylated target samples were used as controls. (**B**) Methylation analysis by COBRA. The restriction enzyme BstUI cleaves d(CGCG) sequences while failing to cut d(TGTG) sites that result from bisulfite conversion. Restriction patterns are shown that were obtained with the 790 bp amplicon of the *p16* promoter. Partial digestion was observed with DNA of HCT116, indicating partial methylation. The PCR-fragment made from SW480 cells was completely digested, which conforms to complete methylation. Analysis of the SW48 material revealed only traces of the bands that result from cleavage and can therefore be considered to be largely unmethylated.

with sodium bisulfite, chemically introducing a C to T conversion at unmethylated cytosines. For MSP, we used two previously published primer pairs (8) that anneal to the positions of cytosines 738, 743, 745 and 854, 858, 860 and 871 (Figure 1). The results indicated that some copies of the *p16* locus were methylated (M) and others unmethylated (U) in HCT116 cells (Figure 2A). For SW480 cells, the MSP resulted in a distinct M band, while for SW48 cells the gel showed a prominent U band in addition to a weak M band (Figure 2A). *In vitro* methylated and unmethylated templates exhibited results that were consistent with the methylation status of the respective templates (Figure 2A). For an independent confirmation of these results, we again deaminated genomic DNA and analysed *p16* methylation by COBRA (9). The resulting PCR fragment was digested with BstUI, which cuts the fragment only if the d(CGCG) target site had been retained through bisulfite-mediated deamination. Enzymatic cleavage thus indicates methylation, while lack of digestion indicates absence of methylation. As expected, we observed a partial digestion of the HCT116 amplicon (Figure 2B), which is in agreement with the existence of both a hyper- and a hypomethylated *p16* allele in HCT116 cells (21). In contrast, the PCR amplicon generated from SW480 cells was completely digested (Figure 2B), which is consistent with dense methylation (22). The SW48 amplicon was largely undigested (Figure 2B) and was therefore considered to be predominantly unmethylated. These results indicated substantial differences

in the *p16* methylation status between the three cell lines and thus established *p16* as a stringent test locus for the establishment of a methylation-sensitive oligonucleotide array.

p16 methylation analysis by oligonucleotide microarrays

After having established the *p16* methylation pattern by conventional methods, we extended our analysis to 53 CpG dinucleotides in the *p16* promoter. For this purpose, we generated oligonucleotide arrays on *Geniom One* DNA processors. For each cell line, a biotinylated *p16* amplicon was hybridized to the first of the four identical oligonucleotide arrays that are present on a single DNA processor. Each array consisted of 7260 oligonucleotide probes, divided into 7008 analysis and 252 control probes, which represent positive controls for the occurrence of DNA synthesis and hybridization or independent sequences for the measurement of background signal. The latter probes reflected the methylation status of every CpG dinucleotide in the sequence of the CpG island of the *p16* promoter. If an oligonucleotide covered more than one CpG site, probes for every possible combination of methylated and unmethylated cytosines were generated. In addition to the cell line sample, amplicons made from *in vitro* methylated or unmethylated DNA were mixed as control samples with a methylation level of 0%, 50% and 100%, respectively, and hybridized individually to arrays 2, 3 and 4. The predicted methylation level of these templates was confirmed by bisulfite sequencing, which revealed 99% methylation of the *in vitro* methylated template and 0% methylation of the unmethylated template (Figure 3A). After washing and staining with fluorescent dye, the signal intensities on all four arrays were

recorded. For each experiment, the background-corrected raw data of array 1 was compared pairwise with the data obtained on arrays 2, 3 and 4. For each cell line, we observed the highest correlation to the sample with the analogous methylation level (Figure 3B). Also, there was a high degree of correlation of the results obtained on either DNA strand (Figure 4A). To determine the discrimination power of the array in greater detail, we also methylated *p16* PCR amplicons with site-specific bacterial methylases (see Materials and Methods for details). Oligonucleotide probes that cover individual target sites were selected and the methylation level was determined by array hybridization, as described above. Probes, which interrogate exactly an individual methylated CpG site, have shown in all cases significantly higher signals in comparison with those probes containing one or more mismatches (e.g. Figure 4B).

Selection of highly discriminative probe molecules

As could be expected, however, the individual oligonucleotide probes differed in their performance. While many permitted a good discrimination between methylated and unmethylated DNA, others failed to provide conclusive data. Most of the latter consisted of sequences that covered more than a single CpG site. For those oligomers, discrimination between fully methylated or completely unmethylated DNA was also good. However, the occurrence of partial methylation in the target sequence decreased the accuracy of the base calling (Figure 4). For the selection of oligonucleotide probes that exhibit a high discriminatory power, a calibration was performed for each probe on the basis of the hybridizations with the control samples of 0, 50 and 100% methylation. After calculating the
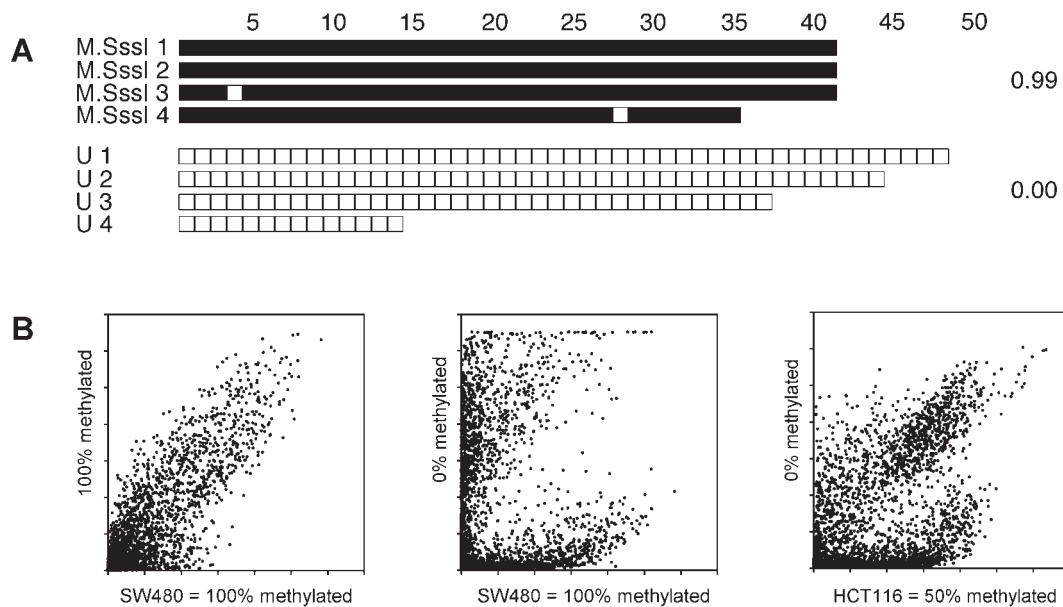

Figure 3. Establishment of a *p16* methylation array. (**A**) Confirmation of the methylation status of *in vitro*-generated *p16* templates by bisulfite sequencing. A sequence of 692 bp of the promoter region including the CpG island is shown. Filled squares indicate methylated cytosines, open squares represent unmethylated cytosines. For the M.SssI-methylated template (M) 99% (0.99) of the CpGs were found to be methylated, while the unmethylated (U) template showed no methylation (0.00). (**B**) Pairwise comparisons of raw hybridization data. Each data point in the scatterplots represents the signal intensities observed at a particular oligonucleotide probe in the hybridizations of a cell line and a control sample. Left panel: amplicon of the cell line SW480 versus the *in vitro* methylated target DNA; the Pearson coefficient is 0.88. Middle panel: cell line SW480 versus unmethylated target DNA; the Pearson coefficient is 0.27. Right panel: cell line HCT116, with 50% *p16* methylation, versus an unmethylated target; a clear separation into signals of high and low correlation can be seen.
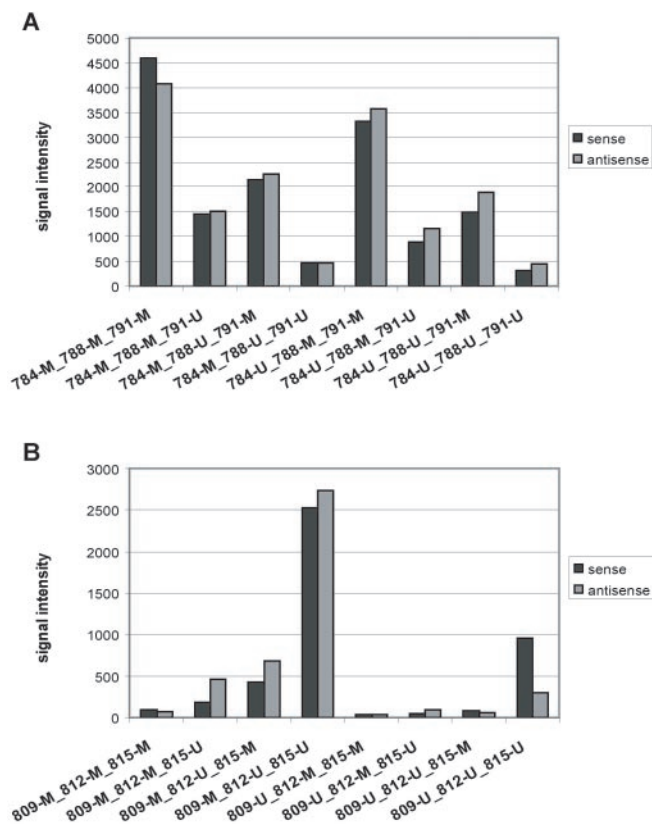
Figure 4. Discrimination power of oligonucleotide probes. (**A**) Hybridization of a fully methylated target. Signal intensities for probes querying the CpG sites 784, 788 and 791 are presented. Signals of the sense probes are shown in grey, the signals on the related antisense probes in black. The signal intensities for the MMM probe (left-most bars) exceed the signals for the UUU probes (right-most bars) by a factor of 11. The signal intensities of the probes interrogating different variations of the methylation status vary considerably depending on the position of the respective mismatch. (**B**) Hybridization of a site-specifically methylated target. While the CpG on position 809 was methylated by using the bacterial methylase M.HpaII, CpG sites 812 and 815 remained unmethylated. The corresponding MUU probes exhibited a prominent, high signal (fourth from the left) which exceeds the signal for the UMM probes (fifth) by a factor of 65 and all other signals by a factor of at least 2.6.

median hybridization signal from eight experimental replications, the intensity ratio M/(M+U) was computed for each oligomer (23). In addition, the Pearson coefficient of correlation of the observed methylation level—measured from the average signal intensities—to the expected methylation level was computed. Only oligonucleotide probes that passed the criteria of producing an intensity ratio of <0.2 for the 0% methylated control target, a ratio of between 0.4 and 0.6 for the 50% methylated target or >0.8 for the 100% methylated target, and additionally exhibited a Pearson coefficient of at least 0.85 were selected for further analyses. More than one half of the examined CpG dinucleotide sequences in one strand were found to be covered by oligonucleotide probes, which met both filtering criteria. Overall, they showed an average intensity ratio of 0.031 for the 0% methylated target, 0.494 for the 50% methylated target and 0.934 for the 100% methylated target (Figure 5). In this highly selective set, the number of oligonucleotides that cover more than a single CpG site increased to 100% from a percentage of 80% in the initial probe set.
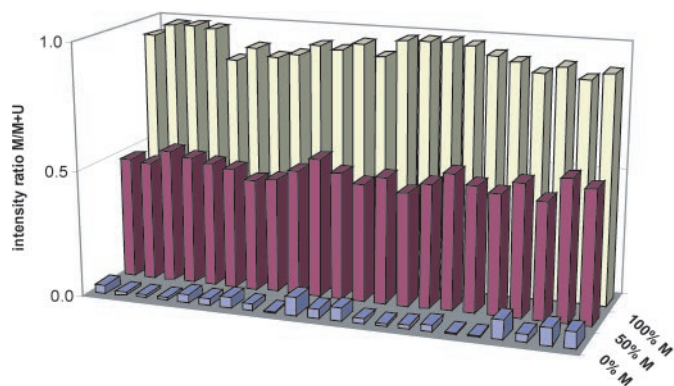

Figure 5. Calibration data of 22 selected oligonucleotide probes. For every oligonucleotide probe a calibration curve was recorded. Using the stringent filter criteria mentioned in the text, probes were selected for the analysis of methylation patterns, which could clearly discriminate among 0, 50 and 100% methylation. For the best 22 probes selected, covering 54% of the CpG positions, the diagram shows the intensity ratios and thus the measured methylation levels obtained with DNA mixtures that represent 0, 50 and 100% methylation. Based on these results, the methylation degree of the studied CpG positions could be accurately quantified.
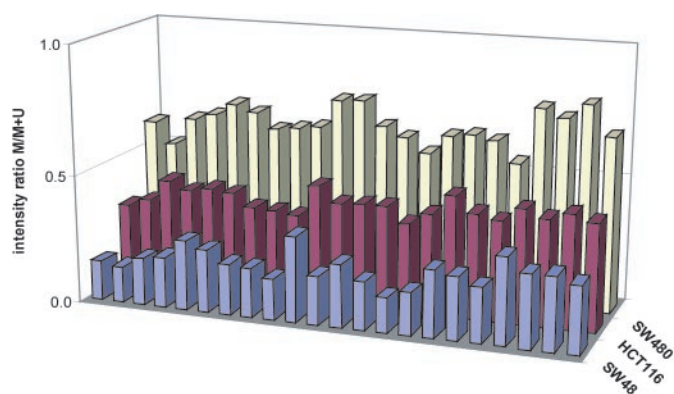

Figure 6. Methylation analysis of the colorectal cancer cell lines. The results of the 22 most discriminative probes are presented. The oligonucleotide microarray assay detected clearly differential methylation between the investigated cell lines SW48, HCT116 and SW480. On average, the *p16* promoter exhibited a methylation level of 20% in SW48 cells, 40% in HCT116 cells and 65% in SW480 cells.

Detailed methylation analysis of genomic DNA from colorectal cancer cell lines

Subsequently, the selected probes were used to analyse the methylation level of the *p16* promoter of the three colorectal cell lines (Figure 6). In agreement with our previous results, they exhibited a clear distinction at the methylation level. Based on the microarray results, the CpG sites of the *p16* amplicon from SW48 cells had a methylation level of 22% (indicated by a mean intensity ratio of 0.220). For HCT116 cells, the methylation level was found to be 40% (mean intensity ratio 0.397). The amplicon of SW480 cells, finally, exhibited a mean intensity ratio of 0.653, corresponding to a methylation level of ~65%.

To confirm the detailed methylation analysis by an independent method, we subjected genomic DNA from HCT116, SW480 and SW48 cell lines to bisulfite sequencing (Figure 7). As predicted by the microarray analysis, the methylation status
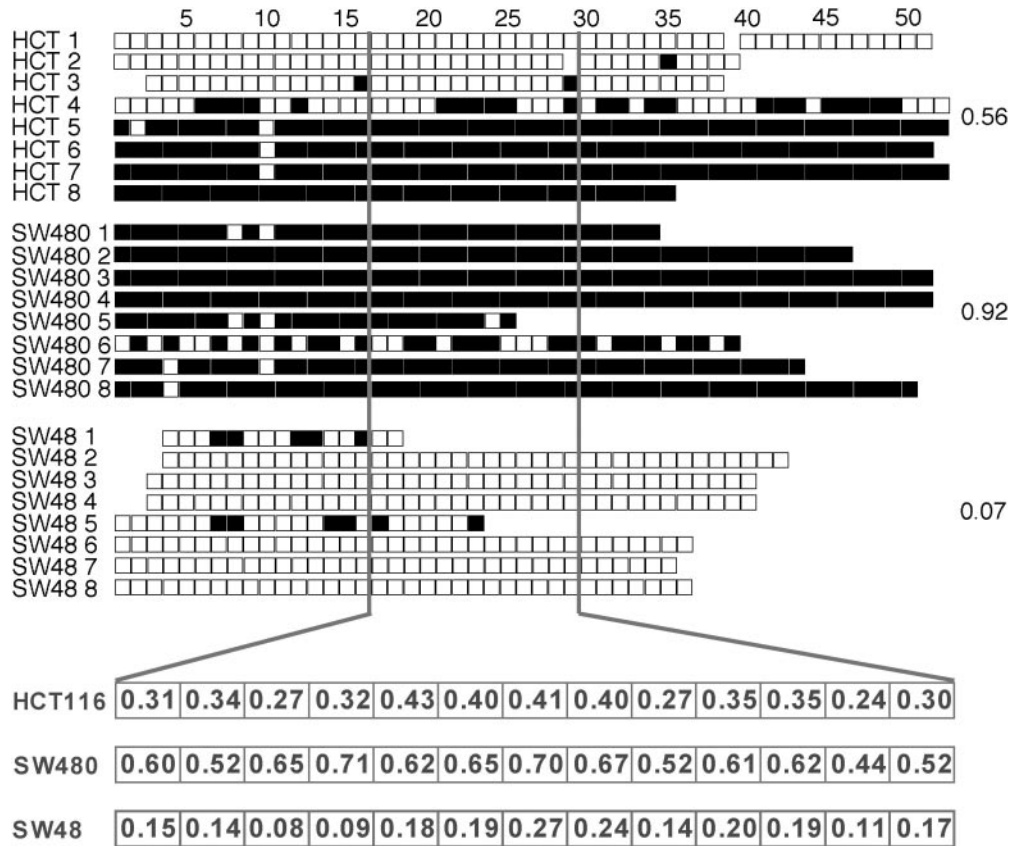
Figure 7. Validation of cytosine methylation patterns by genomic bisulfite sequencing. Genomic DNA was subjected to bisulfite sequencing. The sequences were aligned to the corresponding region of the genomic sequence of *p16* and the methylation status of all cytosine residues was determined. A sequence of 692 bp of the promoter region including the CpG island is shown. Filled squares indicate methylated cytosines, open squares represent unmethylated cytosines. The average methylation level is presented on the right. Below, the average results derived from array-based analyses are shown for a representative region.

at the individual cytosines of the CpG island of the *p16* gene showed clear differences between the colorectal cancer cell lines. HCT116 clearly exhibited a monoallelic methylation. SW480 showed 92% methylation and SW48 only contained ~7% methylated cytosines.

DISCUSSION

We have described an oligonucleotide-based microarray approach for a detailed analysis of the methylation status of many individual CpG sites. Focusing on a CpG island in the promoter of the tumour suppressor gene $p16^{INK4A}$, we demonstrated that this technique provides significant advantages over previous array approaches. By using *in situ* synthesized oligonucleotide microarrays made by the flexible *Geniom One* system, we were able to query individual CpG dinucleotides. The inclusion of probes not only for the sense, but also for the antisense strand of the DNA enlarged the number of accessible CpG positions within the hybridization assay. Nevertheless, there were differences in the performance of individual oligonucleotides. The sequence composition of CpG islands has a low complexity, which may compromise the specificity of oligonucleotide probes. Consequently, the selection of probes suitable for the analysis of DNA methylation patterns is a crucial step. Owing to the very different sequence composition

of the oligonucleotide probes for the methylated and the unmethylated status, the hybridization behaviour with respect to melting temperature and hybridization kinetics may differ a lot, especially for probes containing two or more CpG sites. Cross-hybridization of a methylated target to probes specific for the unmethylated status is more likely than *vice versa*. By applying stringent filter criteria, we were able to select oligonucleotide probes, which allow a reliable and reproducible analysis of the methylation level of many individual CpG sites. For sequences, which are more difficult to be analysed, individual oligomer probes are usually insufficient for a proper discrimination. In such cases, a tiling path of oligonucleotides is required for producing statistically valid information. Alternatively, dynamic allele-specific hybridization (24) or enzymatic primer extension reactions (25) could improve the analysis.

Consistent with published results (21), HCT116 cells usually had a monoallelic methylation of the *p16* promoter region. Most of the sequenced clones were either unmethylated or completely methylated across the entire region that was analysed. However, some incompletely methylated alleles were also observed. This result is in good agreement with an overall methylation level of 40% observed with our microarray experiments. The lower methylation of CpG site no. 10 in Figure 7, as seen in the bisulfite data, was also found in the microarray data, which provided an intensity ratio of 0.115 for

this position, in contrast to an intensity ratio of ∼0.6 of the adjacent CpG dinucleotides.

Our data also suggests that the *p16* promoter is largely unmethylated in SW48 cells. The microarray data, which resulted in a methylation level of 20%, were validated not only by MSP, which yielded a weak M band in addition to the prominent U band, but also by COBRA, which produced some faint bands of digested, therefore methylated DNA besides the distinctive U band. In addition, the bisulfite sequencing also showed a low level of methylation. The slightly lower degree of methylation detected by MSP and COBRA compared to the oligonucleotide microarray is probably due to the fact that methylation analysis by MSP and COBRA is restricted to a few cytosines in the $3'$ end of the primers. Consistent with this notion, a closer analysis of our microarray data revealed an intensity ratio of only 0.09 for the particular CpG sites queried with both MSP and COBRA primers.

In agreement with published data (22), the CpG island of *p16* exhibited a clear hypermethylation in SW480 cells. While MSP and COBRA as well as bisulfite sequencing showed a high degree of methylation, the oligonucleotide microarray displayed a methylation level of 65%. Again, both MSP and COBRA analysed only a limited number of CpG sites, which may not be representative for the whole CpG island. The basic accuracy of the microarray assay had been confirmed in the calibration with the *in vitro* methylated samples. Therefore, it is likely that the difference was introduced during sample processing and preparation. The hybridization assay, for example, directly analyses the mixture of molecules that yield from the bisulfite conversion. In contrast, individual molecules are being analysed by sequencing. Only a limited degree of redundancy is achieved therefore. Adding one or two of the more mosaic clones, such as clone SW480.6 in Figure 7, would change the average methylation value significantly. In addition, the sequenced fragments underwent cloning into *E.coli* prior to the analysis. However, also the hybridization assay might be biased by experimental factors. Since sequences of alternating purines and pyrimidines have a profound effect on the DNA structure, the sequence composition of the CpG island is likely to influence the hybridization efficiency.

In the analysis, several levels of variation were observed. The overall methylation was clearly different between the three colorectal cancer cell lines. However, also significant variations between individual cells of the same type were found at the level of individual cytosines. Additionally, it is apparent that particular cytosines within a CpG island can have a distinct likelihood of being methylated. In all cell lines, the cytosine no. 10 in Figure 7, for example, was significantly less frequently methylated than its neighbours. Whether this is an effect of biological importance and how it is regulated remains to be seen, but this will require large-scale analyses of single-base resolution as presented here. Importantly, this high-resolution determination of an individual cytosine being differentially methylated cannot be obtained by methylation-specific PCR and unlikely to be detected by combined bisulfite restriction analysis.

In summary, our array provides a novel tool for high resolution DNA methylation analysis. The potential of this method will be further expanded in future experiments, in which we will include oligonucleotides from additional genes.

By multiplexing, using different dyes, and by reducing the redundancy of each oligonucleotide well below 8-fold, we will be able to analyse a lot more CpG sites per array. Currently, we are in the process of establishing a chip for ∼250 genes, which are correlated to prostate cancer on the level of transcriptional variations. In addition, we intend to look at the promoter region of all classical tumour suppressor genes. This epigenetic array may eventually be used as a tool to monitor therapeutical interventions before, during and after cancer treatment. With DNA being a much more stable molecule than RNA and with methylation variations occurring slower than changes of the transcript level, epigenetic analyses might be superior to other assays, such as transcriptional profiling, since they are less prone to biases artificially introduced during preparative steps. A comprehensive epigenetic array that allows for a parallel analysis of several epigenetically regulated cancer loci (26–28) may provide detailed insights into the dynamic regulation of genomic DNA methylation and represents a valuable diagnostic tool for oncological research.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

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. Ehrlich,M. (2002) DNA methylation in cancer: too much, but also too little. *Oncogene*, **21**, 5400–5413.
3. Egger,G., Liang,G., Aparicio,A. and Jones,P.A. (2004) Epigenetics in human disease and prospects for epigenetic therapy. *Nature*, **429**, 457–463.
4. Baylin,S.B. and Herman,J.G. (2000) DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.*, **16**, 168–174.
5. Warnecke,P.M. and Bestor,T.H. (2000) Cytosine methylation and human cancer. *Curr. Opin. Oncol.*, **12**, 68–73.
6. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
7. Esteller,M., Sanchez-Cespedes,M., Rosell,R., Sidransky,D., Baylin,S.B. and Herman,J.G. (1999) Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.*, **59**, 67–70.
8. 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.
9. Xiong,Z. and Laird,P.W. (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.*, **25**, 2532–2534.
10. 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.
11. Kennedy,G.C., Matsuzaki,H., Dong,S., Liu,W.M., Huang,J., Liu,G., Su,X., Cao,M., Chen,W., Zhang,J. *et al.* (2003) Large-scale genotyping of complex DNA. *Nat. Biotechnol.*, **21**, 1233–1237.

12. Adorjan,P., Distler,J., Lipscher,E., Model,F., Muller,J., Pelet,C., Braun,A., Florl,A.R., Gutig,D., Grabs,G. *et al.* (2002) Tumour class prediction and discovery by microarray-based DNA methylation analysis. *Nucleic Acids Res.*, **30**, e21.

13. Shi,H., Maier,S., Nimmrich,I., Yan,P.S., Caldwell,C.W., Olek,A. and Huang,T.H. (2003) Oligonucleotide-based microarray for DNA methylation analysis: principles and applications. *J. Cell. Biochem.*, **88**, 138–143.

14. Hou,P., Ji,M., Liu,Z., Shen,J., Cheng,L., He,N. and Lu,Z. (2003) A microarray to analyze methylation patterns of p16(Ink4a) gene 5′-CpG islands. *Clin. Biochem.*, **36**, 197–202.

15. Huang,T.H., Perry,M.R. and Laux,D.E. (1999) Methylation profiling of CpG islands in human breast cancer cells. *Hum. Mol. Genet.*, **8**, 459–470.

16. Baum,M., Bielau,S., Rittner,N., Schmid,K., Eggelbusch,K., Dahms,M., Schlauersbach,A., Tahedl,H., Beier,M., Guimil,R. *et al.* (2003) Validation of a novel, fully integrated and flexible microarray benchtop facility for gene expression profiling. *Nucleic Acids Res.*, **31**, e151.

17. Merlo,A., Herman,J.G., Mao,L., Lee,D.J., Gabrielson,E., Burger,P.C., Baylin,S.B. and Sidransky,D. (1995) 5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nature Med.*, **1**, 686–692.

18. Beier,M. and Hoheisel,J.D. (2000) Production by quantitative photolithographic synthesis of individually quality checked DNA microarrays. *Nucleic Acids Res.*, **28**, e11.

19. Hasan,A., Stengele,K.-P., Giegrich,H., Cornwell,P., Isham,K.R., Sachleben,R.A., Pfleiderer,W. and Foote,S. (1997) Photolabile protecting groups for nucleotides: synthesis and photodeprotection rates. *Tetrahedron*, **53**, 4247–4264.

20. Feinberg,A.P. and Vogelstein,B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6–13.

21. Myohanen,S.K., Baylin,S.B. and Herman,J.G. (1998) Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. *Cancer Res.*, **58**, 591–593.

22. Herman,J.G., Merlo,A., Mao,L., Lapidus,R.G., Issa,J.P., Davidson,N.E., Sidransky,D. and Baylin,S.B. (1995) Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res.*, **55**, 4525–4530.

23. Gitan,R.S., Shi,H., Chen,C.M., Yan,P.S. and Huang,T.H. (2002) Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Res.*, **12**, 158–164.

24. Jobs,M., Howell,W.M., Strömqvist,L., Mayr,T. and Brookes,A.J. (2003) DASH-2: flexible, low-cost, and high-throughput SNP genotyping by dynamic allele-specific hybridization on membrane arrays. *Genome Res.*, **13**, 916–924.

25. Syvanen,A.C. (1999) From gels to chips: 'Minisequencing' primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum. Mutation*, **13**, 1–10.

26. Paz,M.F., Fraga,M.F., Avila,S., Guo,M., Pollan,M., Herman,J.G. and Esteller,M. (2003) A systematic profile of DNA methylation in human cancer cell lines. *Cancer Res.*, **63**, 1114–1121.

27. Qu,G.Z., Grundy,P.E., Narayan,A. and Ehrlich,M. (1999) Frequent hypomethylation in Wilms tumors of pericentromeric DNA in chromosomes 1 and 16. *Cancer Genet. Cytogenet.*, **109**, 34–39.

28. Widschwendter,M., Jiang,G., Woods,C., Muller,H.M., Fiegl,H., Goebel,G., Marth,C., Muller-Holzner,E., Zeimet,A.G., Laird,P.W. and Ehrlich,M. (2004) DNA hypomethylation and ovarian cancer biology. *Cancer Res.*, **64**, 4472–4480.