nature genetics

Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells

Michael Weber[1], Jonathan J Davies[2], David Wittig[1], Edward J Oakeley[1], Michael Haase[3], Wan L Lam[2] & Dirk Schübeler[1]

Cytosine methylation is required for mammalian development and is often perturbed in human cancer. To determine how this epigenetic modification is distributed in the genomes of primary and transformed cells, we used an immunocapturing approach followed by DNA microarray analysis to generate methylation profiles of all human chromosomes at 80-kb resolution and for a large set of CpG islands. In primary cells we identified broad genomic regions of differential methylation with higher levels in gene-rich neighborhoods. Female and male cells had indistinguishable profiles for autosomes but differences on the X chromosome. The inactive X chromosome (Xi) was hypermethylated at only a subset of gene-rich regions and, unexpectedly, overall hypomethylated relative to its active counterpart. The chromosomal methylation profile of transformed cells was similar to that of primary cells. Nevertheless, we detected large genomic segments with hypomethylation in the transformed cell residing in gene-poor areas. Furthermore, analysis of 6,000 CpG islands showed that only a small set of promoters was methylated differentially, suggesting that aberrant methylation of CpG island promoters in malignancy might be less frequent than previously hypothesized.

Reversible methylation of cytosine is a major epigenetic modification in multicellular organisms[1]. In mammals, cytosine methylation occurs almost exclusively at CpG dinucleotides, which are underrepresented in the genome with the exception of CpG islands. These are small CpG-rich regions that, in many cases, are associated with promoter regions. Cytosine methylation results in transcriptional repression either by interfering with transcription factor binding or by inducing a repressive chromatin structure[2]. DNA methylation is required to complete embryonic development[3] and has been directly implicated in genomic imprinting[4] and X-chromosome inactivation[2].

Alterations in DNA methylation are associated with many human diseases and are a hallmark of cancer[5]. A decrease in the total amount of cytosine methylation is observed in many human neoplastic tissues, but the genomic context of this hypomethylation has not been identified[6]. At the same time, aberrant promoter hypermethylation has been observed in sporadic cancer and is thought to contribute to carcinogenesis by inactivating tumor-suppressor genes[5]. In light of the relevance of DNA methylation for normal development and disease, we know little about its genomic distribution. This partly reflects the limitations of existing techniques for analyzing DNA methylation at specific sequences. Here we used an immunocapturing approach to enrich methylated DNA and combine it with detection by DNA microarray. Using whole-genome as well as promoter-specific arrays, we present a methylation profile of unique sequences of the human genome in primary and transformed cells.

RESULTS

Unbiased detection of methylated DNA by immunoprecipitation

Current strategies to identify chromosomal sites of DNA methylation rely primarily on the use of methylation-sensitive restriction enzymes. These require high-molecular-weight DNA and are limited by the sequence context of the chosen enzyme. For example, only 3.9% of all CpGs in human nonrepetitive DNA reside at sites for the frequently used *Hpa*II enzyme[7]. Conversion of unmethylated cytosine with bisulfite followed by sequencing provides an unbiased and sensitive alternative, but it is laborious and cannot be easily applied to screening a large set of sequences or samples[8]. To circumvent these limitations, we developed methylated DNA immunoprecipitation (MeDIP), which permits highly efficient enrichment of methylated DNA. In this assay, an antibody specific for methylated cytosines is used to immunocapture methylated genomic fragments. The resulting enrichment in the immunoprecipitated fraction is determined by standard DNA detection (**Fig. 1a**); thus, MeDIP can be combined with large-scale analysis using existing DNA microarrays.

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. [2]British Columbia Cancer Research Center, Vancouver, British Columbia, Canada. [3]Department of Pathology, Dresden University of Technology, Dresden, Germany. Correspondence should be addressed to D.S. (dirk@fmi.ch).

Published online 10 July 2005; doi:10.1038/ng1598

After optimizing the immunoprecipitation conditions, we carried out a number of control experiments to test further the specificity and efficiency of MeDIP. First, we compared the relative enrichment of known methylated and unmethylated genomic sequences. MeDIP enriched methylated DNA relative to CpG free and unmethylated controls by up to 90-fold (**Fig. 1b**). Next, we analyzed the imprinted *H19* imprinting control region (ICR) sequence, which was previously shown by bisulfite sequencing to be consistently methylated on one allele in all somatic cells analyzed[9,10]. We created defined fragments of the ICR sequence, which contain different numbers of methylated cytosines, by restriction digestion of genomic DNA. The level of enrichment by MeDIP for each sequence increased in a linear manner with the number of methylated cytosines (**Fig. 1c**). The monoallelic methylation of the ICR also allowed us to monitor potential sequence bias. We used a hybrid mouse strain with a polymorphic restriction site in this locus that allowed us to distinguish the parental alleles. Only the methylated paternal allele was detected by MeDIP under our experimental conditions (**Fig. 1d**). Thus, of two alleles with similar sequence, only the methylated one was recognized. These controls indicate that enrichment of 5-methylcytosine by MeDIP occurs in a dose-dependent and sequence-independent manner.

Genomic methylation profiles in male and female primary cells

Because MeDIP allows sensitive detection of cytosine methylation, we combined it with microarray analysis to generate comprehensive maps of DNA methylation of the human genome. We labeled MeDIP-enriched and input genomic DNA with different fluorescent dyes and hybridized them to a microarray. We then calculated the ratio of methylated to input signal for each sequence spotted on the array and used this as a read-out for the methylation level (**Fig. 2a**). A positive log ratio indicates hypermethylation, and a negative log ratio indicates hypomethylation. For these experiments, we used a submegabase-resolution tiling (SMRT) array consisting of 32,433 overlapping BAC clones, which provide 1.5-fold coverage of the human genome[11]. The resulting chromosomal maps represent the methylation landscape of all nonrepetitive genomic sequence, are mostly gap-free and have an average tiling resolution of 80 kb.

Because little information exists regarding the conservation of DNA methylation levels between individuals, we first analyzed genetically unrelated male and female primary nontransformed human fibroblasts. If DNA methylation profiles are conserved, we would expect similar methylation profiles on autosomes between both samples. Methylation levels in male and female cells were almost indistinguishable in this analysis ($R = 0.88$; **Fig. 2b**). We then arranged the BAC clones along their chromosomal locations to create chromosomal methylation profiles, which showed that neighboring sequences tended to have similar methylation levels. We carried out a statistical analysis to test for similar methylation levels of adjacent probes, which confirmed a strong autocorrelation over seven to ten BAC clones and confirmed that methylation tended to be similar over extended

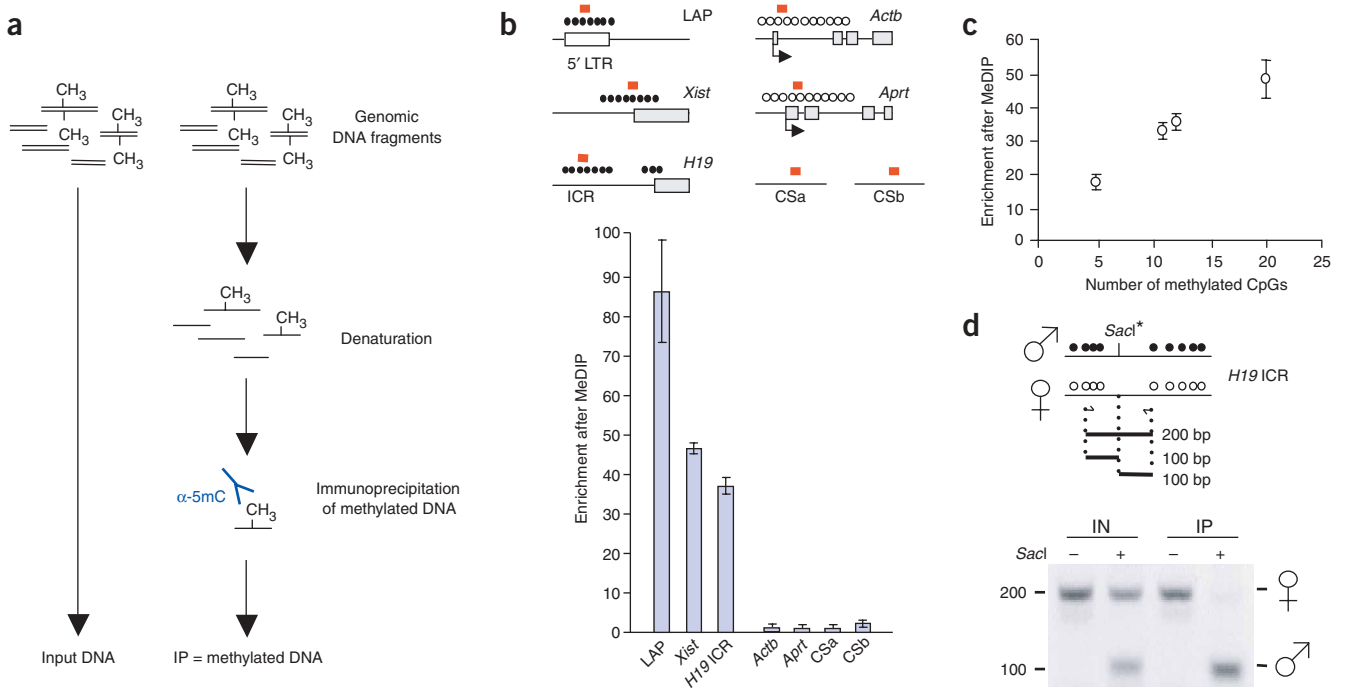

Figure 1 Methylation analysis by DNA immunoprecipitation (MeDIP). (**a**) Denatured genomic DNA of desired fragment length (generated by restriction or sonication) is incubated with an antibody directed against 5-methyl-cytosine (α-5mC), and methylated DNA is isolated by immunoprecipitation (IP). Enrichment of target sequences in the methylated fraction can be quantified by standard DNA detection methods such as PCR or slot blot (*e.g.*, microarray). (**b**) Control sequences that are highly methylated (LAP, *Xist*, *H19*), are unmethylated (*Actb*, *Aprt*) or lack CpGs (CSa, CSb) were selected from the mouse genome. Red bars represent the amplified PCR fragments. MeDIP was done on *Alu*I-treated female genomic DNA, and the relative enrichments in the bound over input fractions was calculated by real-time PCR. The graph shows a specific and efficient enrichment of methylated over unmethylated sequences. (**c**) Correlation between enrichment and the number of methylated cytosines on four *Alu*I restriction fragments in the *H19* ICR sequence. Enrichment increases linearly with increasing number of methylated cytosines. (**d**) Selective immunoprecipitation of the methylated allele at the imprinted *H19* ICR locus. A polymorphic *Sac*I site is used to distinguish the maternal from the paternal allele in a hybrid background (*Mus musculus domesticus* × SD7). Filled and open circles represent methylated and unmethylated CpGs, respectively. Plotted are the average enrichments of three experiments. IN, input; IP, immunoprecipitation.

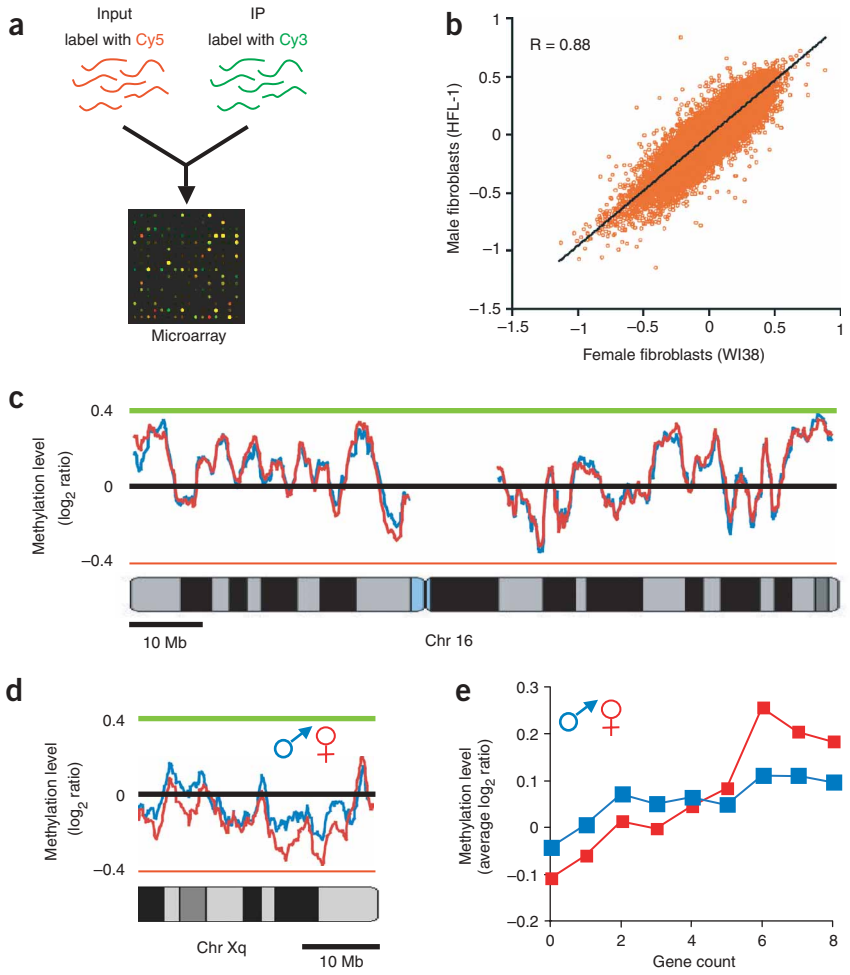

Figure 2 Chromosomal profile of DNA methylation using a tiled whole human genome BAC array. (**a**) Methylation levels are determined by cohybridizing input and bound (IP) DNA labeled with different fluorescent dyes. Methylated sequences were labeled with Cy3 and have a higher fluorescence in the green channel. The ratio of bound over input is calculated for each spot on the array and used as a measure of methylation. (**b**) Comparison of enrichments from genetically unrelated primary fibroblasts. Shown is a pairwise comparison of methylation levels ($\log_2$ ratio) in male and female fibroblasts for all autosomal BACs. The similarity is indicated by a high correlation coefficient (R = 0.88). (**c**) Methylation profile of human chromosome 16 in female (red line) and male (blue line) fibroblasts after local averaging. Gray and black boxes reflect R and G banding, respectively. A methylation map for the complete human genome is shown in **Supplementary Figure 1** online. (**d**) Methylation profile of a 30-Mb region positioned at Xq. The X chromosome is globally hypomethylated in female fibroblasts (red line) compared with male fibroblasts (blue line), except in the gene-rich region at the telomeric end. (**e**) Xi is hypomethylated in gene-poor regions and hypermethylated in gene-rich regions relative to the active X. Plotted is the average methylation level for all X-linked BAC clones depending on gene count for male (blue) and female (red) fibroblasts.

genomic regions (data not shown). We then applied local averaging to create complete chromosomal maps of DNA methylation in male and female fibroblasts (**Fig. 2c** and **Supplementary Fig. 1** online). These profiles illustrate the high similarity of DNA methylation between male and female autosomes.

The inactive X chromosome is globally hypomethylated

Promoter hypermethylation at the inactivated X chromosome (Xi) is a characteristic of mammalian dosage compensation[12]. It is unclear how the global distribution of methylation is affected on Xi. Our data set allowed us to compare methylation profiles in male and female fibroblasts over the entire X chromosome, which is represented by 1,461 BAC clones. The methylation level measured in female cells is the sum of the methylation levels on the active and inactive X chromosomes; therefore, the differences between the two alleles are underestimated in our experimental setup.

A male-female comparison of methylation levels for each chromosome showed that the X chromosome had a marked difference in methylation ($P = 10^{-66}$) compared with the autosomes (**Supplementary Fig. 2** online). The X chromosome in female cells was generally hypomethylated compared with the male X chromosome (**Fig. 2d**). This global hypomethylation does not reflect perturbed X inactivation in the particular cells studied, as PCR analysis detected local hypermethylation of different X-linked promoters in female cells (data not shown). A more detailed analysis showed that gene-rich regions of the X chromosome were more methylated in the female cells (**Fig. 2d,e**), probably reflecting *de novo* methylation of promoters undergoing X inactivation. Yet most of the X chromosome is gene-poor and, according to our experiments, is hypomethylated at Xi. Previous reports had already hinted at reduced methylation of Xi in metaphase spreads, as measured by immunostaining with an antibody against 5-methylcytosine[13] or by immunodetection of unmethylated *Hha*II restriction sites[14]. In addition, there is at least one reported example of a low-copy repeat sequence that is unmethylated exclusively on Xi[15]. Our analysis extends these findings by providing evidence for global Xi hypomethylation.

Genomic determinants of differentially methylated regions

Next, we assessed the sequence characteristics related to differential methylation in the genome. A visual inspection of the chromosomal profiles (**Fig. 2c** and **Supplementary Fig. 1** online) indicates that regions with high methylation levels tend to localize more often in R bands than in G bands. Even though the exact nature of R-G band staining is not known, R bands tend to be gene-rich and have a higher GC content than G bands. Therefore, we collected information on gene count, Alu and LINE content and GC percentage for each probe present on the array and determined how these sequence characteristics related to the measured degree of methylation. On the level of whole chromosomes, gene-rich chromosomes had higher levels of cytosine methylation than gene-poor chromosomes (R = 0.93; **Fig. 3a**). Therefore, when averaged over large genomic regions, methylation correlated perfectly with gene density. Because gene-rich

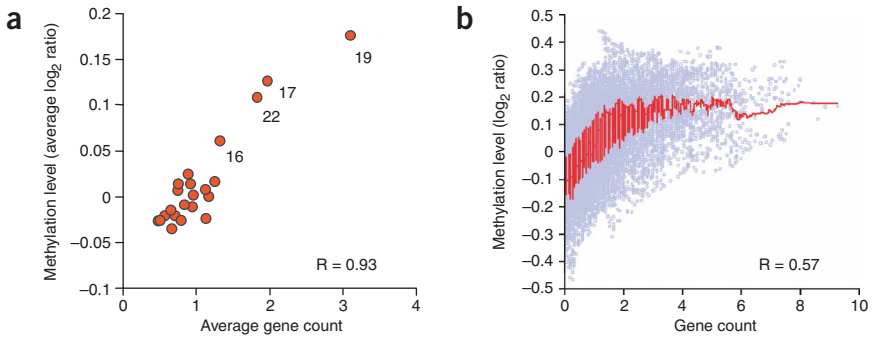

Figure 3 Methylated cytosines are more abundant in gene-rich regions. (**a**) The average methylation level was calculated in primary fibroblasts for each chromosome and plotted against the average gene count. Gene-rich chromosomes have a higher average methylation level, as indicated by a high positive correlation (R = 0.93). (**b**) Reduced correlation between methylation level and gene count at the level of BAC probes. The red line represents a moving average over 50 probes.

domains of the genomes also have high CG and Alu contents[16], there was a similar tight correlation with chromosomal Alu (R = 0.97) and GC content (R = 0.94; **Supplementary Fig. 3** online). For LINE elements, however, which are more abundant in gene-poor regions, we observed a negative correlation (R = −0.21; **Supplementary Fig. 3** online). We conclude that, on the chromosomal level, gene count and GC and Alu content can be excellent predictors of DNA methylation.

At the level of BAC probes, these general trends were still present, but the correlation of DNA methylation with gene density (R = 0.57; **Fig. 3a**) and with Alu (R = 0.56) and GC content (R = 0.38) was strongly reduced (**Fig. 3b** and **Supplementary Fig. 3** online). For any given gene count or GC or Alu content, the methylation level of BACs can vary widely, as suggested by the reduced correlation. Therefore, at the local level compared with the larger chromosomal scale, gene count and sequence composition are rather poor predictors of DNA methylation. We propose that this difference reflects the influence of large surrounding domains on the methylation level at any given chromosomal position.

Chromosomal profiles of DNA methylation in a cancer cell line

Numerous reports have indicated that some tumor cells have an imbalance in DNA methylation[5]. The emerging picture shows a global reduction in the amount of methylated cytosine with coinciding hypermethylation of a subset of promoters, which may be linked to inactive tumor-suppressor genes. We mapped the global distribution of DNA methylation in a colon cancer model (SW48 cells) to approximate the conservation of chromosomal methylation in a transformed cell.

The resulting genomic profile indicates that the global distribution of cytosine methylation is markedly similar to that of primary

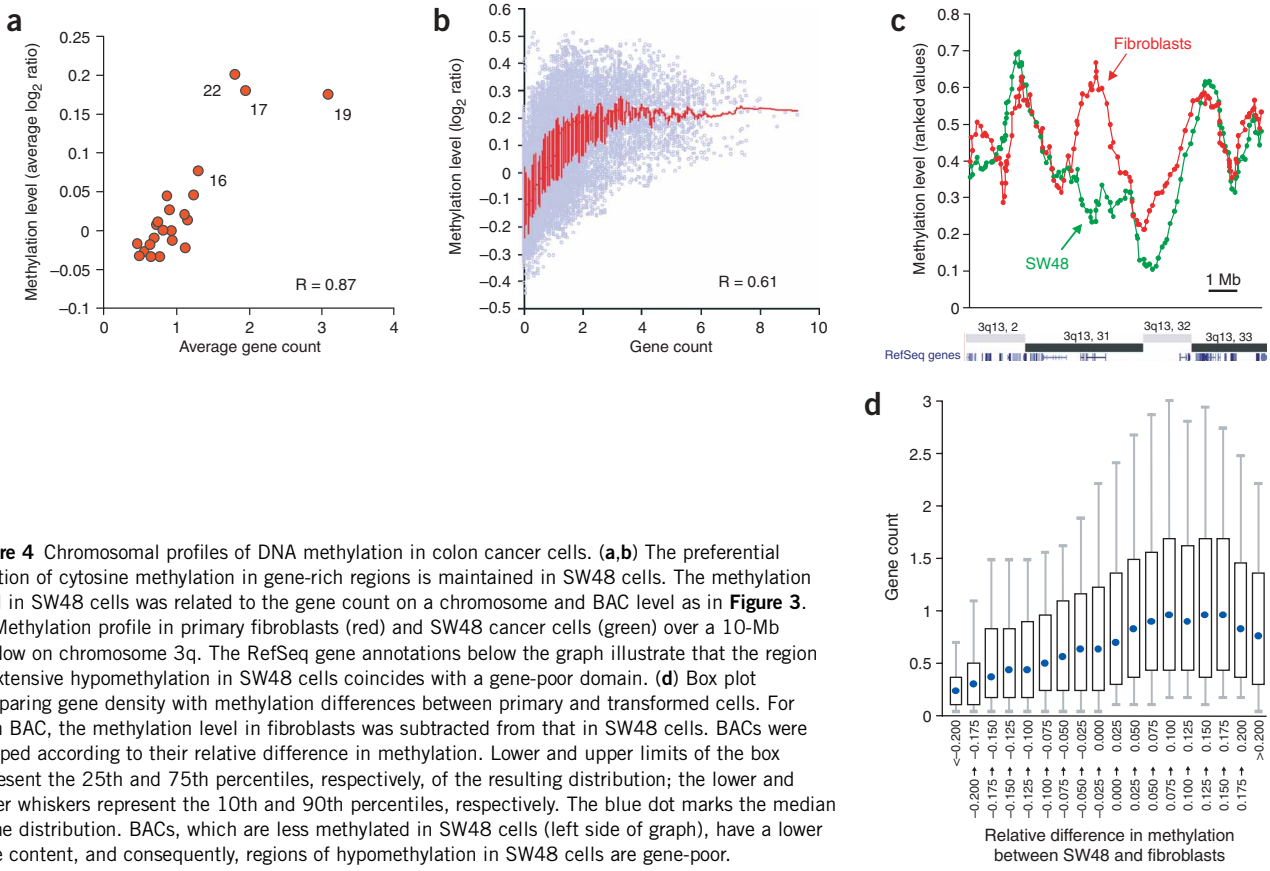

Figure 4 Chromosomal profiles of DNA methylation in colon cancer cells. (**a**,**b**) The preferential location of cytosine methylation in gene-rich regions is maintained in SW48 cells. The methylation level in SW48 cells was related to the gene count on a chromosome and BAC level as in **Figure 3**. (**c**) Methylation profile in primary fibroblasts (red) and SW48 cancer cells (green) over a 10-Mb window on chromosome 3q. The RefSeq gene annotations below the graph illustrate that the region of extensive hypomethylation in SW48 cells coincides with a gene-poor domain. (**d**) Box plot comparing gene density with methylation differences between primary and transformed cells. For each BAC, the methylation level in fibroblasts was subtracted from that in SW48 cells. BACs were grouped according to their relative difference in methylation. Lower and upper limits of the box represent the 25th and 75th percentiles, respectively, of the resulting distribution; the lower and upper whiskers represent the 10th and 90th percentiles, respectively. The blue dot marks the median of the distribution. BACs, which are less methylated in SW48 cells (left side of graph), have a lower gene content, and consequently, regions of hypomethylation in SW48 cells are gene-poor.

fibroblasts (**Supplementary Fig. 4** online), as indicated by a high pairwise correlation of the methylation levels ($\log_2$ ratios) between both genomes ($R = 0.61$). In particular, the preferential location of DNA methylation in gene-rich regions was highly evident in SW48 cells. Similar to the results obtained in primary cells, we found the highest levels of methylation at gene-rich chromosomes and gene-rich probes (**Fig. 4a**,**b**).

We also detected several regions (up to 20 Mb in size) with marked hypomethylation in SW48 cells, including 3p26, 3q13.31, 7q35, 14q31 and 11q22.3 (**Fig. 4c** and **Supplementary Fig. 4** online). We carried out a global comparative genomic hybridization analysis[11] to determine whether local methylation changes reflect duplication or deletion events. This analysis showed that primary cells were of normal diploid karyotype, whereas SW48 cells were trisomic with respect to chromosomes 7 and 14 and a subregion in chromosome 10q (**Supplementary Fig. 4** online). Because most regions of differential methylation do not localize to sites of amplification or deletion events, aneuploidy is not required for the observed differences. But chromosomes 7 and 14 are among those with high overall hypomethylation in SW48 cells; therefore, hypomethylation could have contributed to chromosomal instability. To test further whether hypomethylation in SW48 cells was linked to gene density, we calculated the actual differences in DNA methylation between both cell types for any genomic probe and compared it with its gene content (**Fig. 4d**). This analysis indicated that SW48-specific hypomethylation occurred almost exclusively in gene-poor neighborhoods of the genome. Additional analysis of multiple tumor samples with matched controls will be required to elucidate whether these differences in methylation can be linked to cancer progression.

CpG island methylation in normal and transformed cells

Having identified extended chromosomal regions of differential methylation, we next asked how they related to the epigenetic state of individual CpG islands. We combined MeDIP with hybridization to a microarray representing ~12,000 CpG island probes. The sequences present on this array are derived from a CpG island library in which 75% of all clones represent unique sequences[17] and 25% correspond to repetitive elements as well as ribosomal and mitochondrial DNA. Half of the unique clones map to the 5′ end of known genes and thus represent *bona fide* promoters. Hybridization resulted in high-quality measurements for 6,000 sequences, and repeat experiments showed high reproducibility.

In primary female fibroblasts, most of the CpG islands had only a basal enrichment in the immunoprecipitated fraction, suggesting that they remain unmethylated. A subset of sequences, corresponding mostly to mitochondrial DNA, which is known to lack cytosine methylation[18], had an even lower level of enrichment (**Fig. 5**). Among the sequences with high levels of methylation, we identified a majority with similarity to repetitive DNA (satellites, LINEs and SINEs), as well as promoters of genes that are imprinted or reside on the X chromosome. On average, X-linked promoters had a much higher enrichment than autosomal clones ($P < 0.001$), reflecting promoter hypermethylation during X inactivation. We conclude that our experimental strategy is suitable for identifying methylated CpG islands and, in agreement with current models[1], that CpG islands remain predominantly unmethylated in primary cells.

To assess the frequency of aberrant CpG island methylation in transformed cells, we generated a CpG island methylation profile in SW48 colon cancer cells and compared it with those of primary fibroblasts and normal colon mucosa. This analysis showed that methylation levels of most CpG islands were maintained in the SW48 cancer cell line (**Fig. 5**). We did not detect sequences that were hypomethylated only in SW48 cells, suggesting that, in this experimental system, epigenetic misregulation does not involve frequent demethylation of CpG islands. But we did identify clones that showed hypermethylation exclusively in the colon cancer line. By comparing SW48 cells with primary fibroblasts, we identified 210 clones that were hypermethylated only in SW48 cells, 112 of which could be identified unambiguously on the basis of sequence annotation. Of these, only 32 correspond to unique sequences and 80 represent ribosomal DNA. Such hypermethylation of ribosomal DNA has previously been reported in relation to aging and neoplasia, but its physiological role is not known[19,20]. Of the unique sequences, 4 clones represent intergenic CpG islands, and the remaining 28 map to 26 different genes. With the exception of two genes, all were located in the promoter region. Notably, when comparing SW48 cells with normal colon mucosa, we identified an almost identical population of unique sequences (**Fig. 5b**), indicating that this methylation was linked to the transformed state and did not represent colon-specific promoter methylation.

To validate the microarray results by single-gene PCR on methylated DNA, we designed primers specific for 22 of these

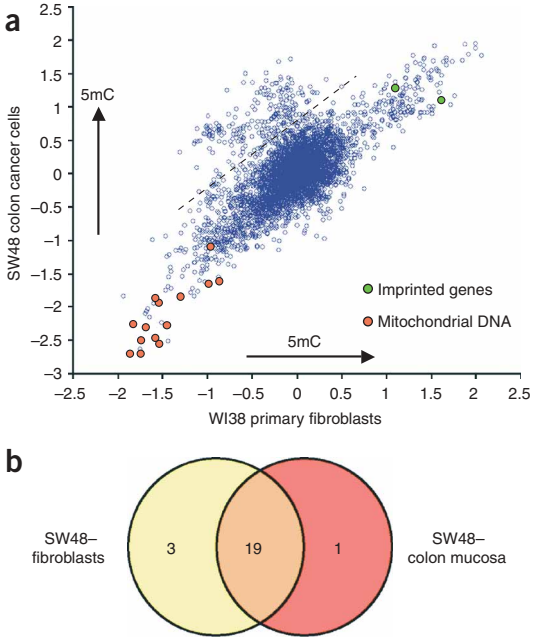

Figure 5 CpG island methylation profile in colon cancer cells versus primary fibroblasts and normal colon mucosa. (**a**) Methylated DNA was analyzed on a microarray representing CpG islands of the human genome. Shown is a pairwise comparison of methylation levels ($\log_2$ ratios) between SW48 colon cancer cells and WI38 primary fibroblasts for 5,000 clones. CpG islands corresponding to the promoter of known imprinted genes or to mitochondrial DNA are marked in green or red, respectively. The dashed line marks the threshold cut-off (residual > 0.75) that was applied to isolate hypermethylated CpG islands in cancer cells. Values represent the average of two independent experiments. (**b**) CpG island methylation in SW48 cells was compared with that of primary fibroblasts and normal colon mucosa. The Venn diagram compares the populations of unique sequences identified as hypermethylated in SW48 cells in both analyses. The high overlap indicates that the two populations are nearly identical.

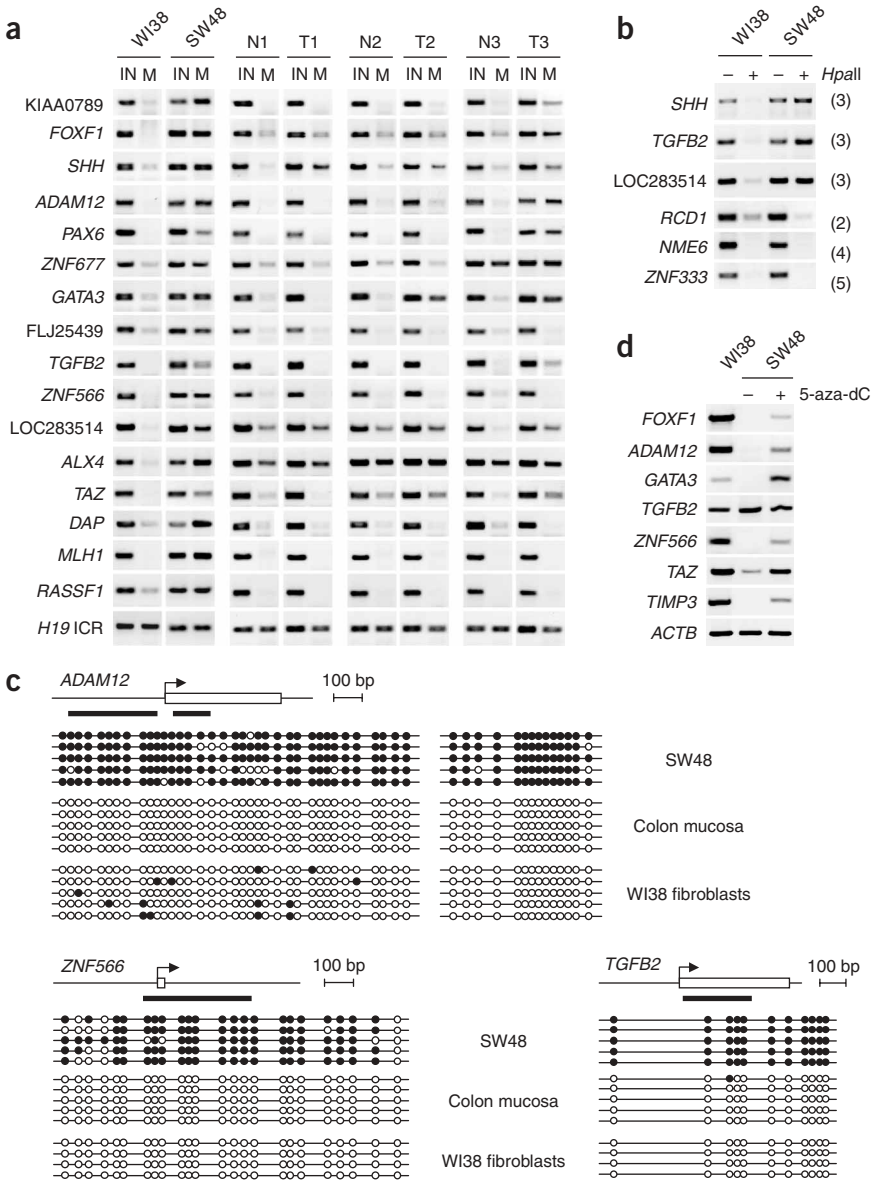

Figure 6 New targets for aberrant methylation in colon cancer. (**a**) Validation of clones identified as hypermethylated in SW48 cells in microarray analysis by single-gene PCR. DNA methylation was controlled by single-gene PCR on MeDIP samples prepared from female primary fibroblasts, SW48 colon cancer cells and matched normal colon (N) and colon tumor (T) from three individuals. IN, input genomic DNA; M, MeDIP enriched methylated DNA. *MLH1* and *RASSF1* were previously described to be aberrantly methylated in SW48 cells and in some colon tumors. The imprinted *H19* ICR sequence serves as a positive control for methylation. (**b**) Control of methylation status by methylation-sensitive restriction. Genomic DNA was either digested with the methylation-sensitive *Hpa*II enzyme or undigested and used as a PCR template with primers spanning a *Hpa*II-containing PCR fragment in three randomly selected positive clones (top) and negative clones (bottom). The number of *Hpa*II sites in the PCR amplicon is indicated in parenthesis. Presence of a PCR product after *Hpa*II digest reflects DNA methylation in the sample and, in each case, confirms the MeDIP analysis. (**c**) Methylation analysis in the promoter of candidate genes by bisulfite genomic sequencing. For each gene, the transcription start site (arrow) and first exon (open box) are shown, and regions analyzed by bisulfite sequencing are indicated below the gene. Each line represents the sequence of a single clone. CpGs are shown as white (unmethylated) or black (methylated) circles. (**d**) Reactivation of silenced genes in SW48 cells by treatment with 5-aza-dC. RT-PCR was done on cDNA prepared from female fibroblasts and SW48 cells and treated with 5 μM 5-aza-dC for 4 d (+) or untreated (−). In this analysis, *TGFB2* transcripts are also detected in SW48 cells. *ACTB* served as an unmethylated control.

candidate genes. These controls confirmed SW48-specific methylation in 70% of all genes (**Fig. 6a** and **Table 1**). We used methylation-sensitive restriction digestion as a separate independent approach and verified differential methylation on three randomly selected genes (**Fig. 6b**). In addition, we carried out bisulfite genomic sequencing of four genes. In each case, the bisulfite sequencing confirmed the results obtained by MeDIP and showed extensive methylation in SW48 cells but not in normal colon mucosa or primary fibroblasts (**Fig. 6c** and data not shown).

Finally, using RT-PCR analysis of a subset of these genes, we found transcriptional downregulation in SW48 cells and derepression by treatment with the demethylating agent 5-aza-2′-deoxycytidine (5-aza-dC; **Fig. 6d**). This reactivation suggests that the detected methylation is directly repressing the activity of the linked gene. We conclude that combining MeDIP with hybridization on a CpG island microarray allows the identification of epigenetically silenced genes in cancer cells. The resulting global pattern of CpG island methylation is conserved between primary and transformed cells, and the number of hypermethylated CpG island promoters in transformed cells seems to be unexpectedly low.

New targets of aberrant methylation in colon cancer

Among the target genes identified on the CpG island array, only the homeobox gene *PAX6* had already been reported to be methylated in SW48 cells[21]. Other previously described targets such as *MLH1*, *RASSF1*, *TIMP3* or *SLIT2* were not present on the array but could be confirmed by single-gene PCR on MeDIP-enriched DNA (**Fig. 6a** and data not shown). The gene *GATA3*, not studied previously in colon cancer, was reported to be aberrantly methylated in breast cancer cells[22]. The remaining genes are new targets for aberrant hypermethylation in cancer.

These genes are involved in a wide variety of biological functions that include regulation of transcription (*FOXF1*, *PAX6*, *TAZ*, *GATA3*), cell cycle progression (*TGFB2*), cell-matrix interactions (*ADAM12*) and apoptosis (*DAP*). Some (*RASL11A*, *FOXF1*, *TGFB2*) have already been implicated in cancer and have been reported to be down-regulated in some tumors[23–25].

To assess the potential relevance of our findings for cancer biology, we determined the methylation status for this set of genes in primary

Table 1 Genes identified as hypermethylated in SW48 cancer cells

Gene name	Accession number	Location	Residual (log_2 ratio) SW48-WI38	Residual (log_2 ratio) SW48-colon	CpG island
KIAA0789*	NM_014653	12q24.11	1.886	1.053	Promoter
Forkhead box F1 (*FOXF1*)*	NM_001451	16q24	1.584	0.934	Exon 2
Sonic hedgehog homolog (*SHH*)*	NM_000193	7q36	1.310	1.208	Exon 2
Disintegrin and metalloproteinase domain 12 (*ADAM12*)*	NM_003474	10q26.3	1.535	0.957	Promoter
RAS-like, family 11, member A (*RASL11A*)	NM_206827	13q22.2	1.629	0.849	Promoter
Paired box gene 6 (*PAX6*)*	NM_000280	11p13	1.245	1.014	Promoter
Predicted gene	BC038214	1q31.3	1.006	1.140	Promoter
Zinc finger protein 677 (*ZNF677*)*	NM_182609	19q13.42	1.426	0.753	Promoter
GATA binding protein 3 (*GATA3*)*	NM_001002295	10p15	1.109	0.928	Promoter
FLJ25439*	NM_144725	5p13.2	1.217	0.756	Promoter
Cell division cycle associated 2 (*CDCA2*)	NM_152562	8p21.2	0.870	1.099	Promoter
Transforming growth factor, beta 2 (*TGFB2*)*	NM_003238	1q41	1.101	0.811	Promoter
Zinc finger protein 566 (*ZNF566*)*	NM_032838	19q13.13	1.066	0.831	Promoter
Ribosomal protein S27-like (*RPS27L*)	NM_015920	15q22.2	0.988	0.891	Promoter
LOC283514*	NM_198849	13q14.13	0.823	0.907	Promoter
Aristaless-like homeobox 4 (*ALX4*)*	NM_021926	11p11.2	1.040	0.733	Promoter
Transcriptional co-activator with PDZ-binding motif (*TAZ*)*	NM_015472	3q23–q24	1.117	0.417	Promoter
Death-associated protein (*DAP*)*	NM_004394	5p15.2	0.801	0.566	Promoter

The residual value describes the difference in methylation enrichment as measured on the CpG island array. It is calculated by subtracting the log_2 ratio in SW48 cells from that in fibroblasts or normal colon mucosa. The last column indicates the position of the CpG island clone in the gene. Targets highlighted with an asterisk (*) were confirmed by single-gene PCR analysis (**Fig. 6a**).

adenocarcinoma and matched normal colon tissue from three individuals (**Fig. 6a**). Many genes were methylated in one or two of the tested adenocarcinoma but not in the matched normal control. The variable methylation of these genes in the tested adenocarcinoma and normal colon samples highlights the epigenetic complexity in colon cancer. We also detect different levels of existing methylation in normal colon for some genes (*ALX4*, *ZNF677*, LOC283514). Such differential accumulation of methylation with age has been observed for other genes and might predispose to cancer formation[26,27]. In summary, more than half of the tested genes identified as methylated in SW48 cells were also methylated in at least one of three tested tumors, indicating that we identified new targets for aberrant hypermethylation *in vivo*.

DISCUSSION

This study shows that immunocapturing with an antibody against 5-methyl-cytidine after random fragmentation allows highly specific isolation of methylated DNA. This technique circumvents the sequence bias of approaches that rely on restriction digestion[7]. We show that MeDIP can be combined with genomic analysis using existing microarray platforms and does not require particular array probes as, for example, with the analysis of bisulfite-converted DNA[28]. Thus, epigenomic profiling of DNA methylation can be done in existing laboratory settings. Of note, the fragmented DNA used in this study has a size range similar to DNA after isolation from formalin-fixed tissues. Accordingly, it should be possible to use MeDIP to screen stored clinical samples and enable large-scale analysis of material from individuals with defined clinical history.

Our analysis results in a high-resolution analysis of DNA methylation of unique sequences along the human genome and shows that gene-rich domains contain high levels of DNA methylation. These blocks are contiguous over large regions and extend over several genes. A preferential methylation of gene-rich R bands had previously been suggested on the basis of immunostaining of metaphase chromosomes[13,29]. Chromosomal stainings have an approximate resolution of 10 Mb and do not distinguish between unique and repetitive DNA. Our analysis provides a resolution more than 100 times higher under hybridization conditions that largely block repetitive DNA. Therefore, we conclude that the measured methylation must reside mostly in unique or low-copy sequences. Furthermore, it does not reflect methylation of promoter CpG islands as we find these largely unmethylated (**Fig. 5**). On the basis of these findings, we propose that genic regions in general are highly methylated, which is in line with previous single-gene analysis[30].

DNA methylation leads to a repressive chromatin structure through recruitment of histone deacetylase activity[31]. Does intragenic methylation hinder polymerase elongation, as has been shown in the fungus *Neurospora crassa*[32]? A recent study using murine cell lines indicates reduced polymerase elongation when only intragenic DNA methylation was increased at a defined genomic locus[33]. Such intragenic methylation with its potential negative effect on polymerase elongation could inhibit inappropriate transcriptional initiation at cryptic sites[34] and might spread from methylated repeats[35].

Genomic studies using microarrays with BAC-sized probes have recently shown that gene-rich regions of the human genome replicate early during S phase[36,37], reside in open chromatin fibers[38] and localize outside their chromosomal territory in the interphase nucleus[38]. Hence, the same regions that share these euchromatic features contain high levels of DNA methylation. This apparent paradox argues that the local repressive chromatin structure mediated by DNA methylation does not interfere with early replication timing, euchromatic fiber organization or nuclear localization. In this context, cytosine methylation seems to be an epigenetic mark that restricts access to DNA only locally and does not necessarily lead to heterochromatic structures.

Notably, Xi as a form of facultative heterochromatin shows overall reduced methylation with the exception of gene-rich regions. This observation challenges the view that chromosome-wide hypermethylation is a characteristic of X inactivation, and it will be interesting

to determine how the observed hypomethylation relates to the nonuniform chromatin structure of Xi[39,40] and if it is mechanistically involved in dosage compensation. Regardless, the differential abundance of DNA methylation in gene-rich regions of the genome and at X-linked facultative heterochromatin point to a context-dependent function of DNA methylation in euchromatic and heterochromatic regions of the genome.

Our chromosomal and promoter-specific methylation profiles allowed us to approximate the extent and localization of differential methylation between a primary and a transformed cell. The chromosomal patterns in both cell types are markedly similar, suggesting that global patterns of DNA methylation are largely conserved between a colon cancer cell line and primary cells (**Supplementary Fig. 4** online). But we detected defined chromosomal regions of differential methylation, extend over 20-Mb domains (**Fig. 4** and **Supplementary Fig. 4** online). Detailed analysis of these regions showed that specific hypomethylation occurs preferentially in gene-poor regions of the genome of SW48 cells. Further genomic and epigenomic analysis, similar to that reported here, of multiple tumor samples will be required to elucidate whether these extended regions are frequent targets and whether they coincide with sites of genomic instability, as has been suggested on the basis of analysis of knockouts of genes encoding DNA methyltransferases[41–43].

We did not detect preferential localization of aberrantly methylated CpG island promoters in chromosomal regions with differential methylation. This could suggest that these processes are not coordinated, but a more comprehensive analysis including non-CpG island promoters might be required to answer this question conclusively.

We found only a small set of genes to be hypermethylated in SW48 cells compared with primary fibroblasts and normal colon mucosa. A simple extrapolation of our CpG island screen to the remainder of the genome would predict that ~200 unique genes are hypermethylated specifically in SW48 cells, considerably fewer than previously estimated for colon cancer[44]. A small number of epigenetically silenced genes might imply that aberrant hypermethylation is either a random event with low frequency or very selective. Selectivity could either be mediated by combined targeting of coregulated genes (upstream) or be a consequence of clonal selection during neoplasia (downstream). Either case would predict that these genes are transcriptionally repressed as a consequence of hypermethylation and that they are frequently methylated in primary cancer samples. Inhibition of DNA methylation results in transcriptional activation of most of the genes identified in this screen, which we also identified as targets of aberrant promoter methylation in primary adenocarcinoma.

Recent genome-wide studies have shed light on transcriptional regulation[45,46] and the interplay of transcription with genome stability[11,47], chromatin structure and organization[38,48] and DNA replication[36,49]. Our comprehensive analysis provides a first epigenomic map of DNA methylation in the human genome. Additional studies using a similar strategy should yield further insights into the dynamics and hierarchy of epigenetic regulation during normal development and disease.

METHODS

Cell culture and tissue samples. We obtained human primary lung fibroblasts (male, HFL-1; female, WI38) and the colon cancer cell line SW48 from ATCC and cultured them in Dulbecco's modified Eagle medium containing 10% fetal calf serum at 37 °C and 5% CO_2 as described[37]. The fibroblasts represent primary cells as they are nontransformed, are nonclonal and undergo senescence after a limited number of passages, similar to mouse embryonic fibroblasts. We took samples from low–passage number fibroblasts before senescence. Global comparative genomic hybridization analysis showed that these fibroblasts have a perfect karyotype (data not shown). We took adenocarcinoma samples and matched controls from three individuals and immediately froze them. The purity of the tumor samples was 80%, based on standard histology.

5-aza-dC treatment and RT-PCR. We seeded SW48 cells (1×10^6) in culture medium and maintained them for 24 h before treating them with 5 μM 5-aza-dC (Sigma) for 4 d. We renewed medium containing 5-aza-dC every 24 h during the treatment. We handled control cells the same way, without adding 5-aza-dC. We prepared total RNA using the RNeasy Mini Kit (Qiagen) and synthesized cDNA from 2 μg of total RNA using the Superscript first-strand synthesis system (Invitrogen) and oligo-dT primers. We carried out PCR reactions on 1/20 of the cDNA preparation. Controls without reverse transcriptase enzyme were negative. Primer sequences are given in **Supplementary Table 1** online.

MeDIP assay. We prepared genomic DNA from cultured cells and tissue samples by overnight Proteinase K treatment, phenol-chloroform extraction, ethanol precipitation and RNase digestion. Before carrying out MeDIP, we sonicated genomic DNA to produce random fragments ranging in size from 300 to 1,000 bp. If indicated, genomic DNA was digested with *Alu*I to create defined fragments. We used 4 μg of fragmented DNA for a standard MeDIP assay. We denatured the DNA for 10 min at 95 °C and immunoprecipitated it for 2 h at 4 °C with 10 μl of monoclonal antibody against 5-methylcytidine (Eurogentec[29]) in a final volume of 500 μl IP buffer (10 mM sodium phosphate (pH 7.0), 140 mM NaCl, 0.05% Triton X-100). We incubated the mixture with 30 μl of Dynabeads with M-280 sheep antibody to mouse IgG (Dynal Biotech) for 2 h at 4 °C and washed it three times with 700 μl of IP buffer. We then treated the beads with proteinase K for 3 h at 50 °C and recovered the methylated DNA by phenol-chloroform extraction followed by ethanol precipitation.

PCR and real-time PCR on MeDIP samples. We carried out PCR and real-time PCR reactions with 25 ng of input DNA and 1/30 of the immunoprecipitated methylated DNA. For real-time PCR reactions, we used the SYBR Green PCR master mix (Applied Biosystems) and an ABI Prism 7000 Sequence Detection System. Reactions were done in duplicates and standard curves were calculated on serial dilutions (100–0.1 ng) of input genomic DNA. To evaluate the relative enrichment of target sequences after MeDIP, we calculated the ratios of the signals in the immunoprecipitated DNA versus input DNA. The resulting values were standardized against the unmethylated control sequence CSa, which was given the value 1. In case of the regular PCR, the reaction was initially set up on serial DNA dilutions with varying number of cycles to ensure that the PCR amplification is in the linear range. Primer sequences are given in **Supplementary Table 1** online.

SMRT array hybridization and analysis. For the genome-wide profiles of DNA methylation, we used the previously described SMRT array consisting of 32,433 overlapping BAC clones with an approximate resolution of 80 kb (*i.e.*, two-thirds of an average BAC clone)[11]. We spotted the entire set of clones in triplicate onto two aldehyde-coated slides. We labeled 400 ng of sonicated input DNA and of methylated DNA enriched by the MeDIP assay separately with cyanine-3 and cyanine-5 dCTPs. We carried out probe labeling, repeat blocking with Cot-1 DNA, subsequent hybridization and washing as described[11]. We imaged hybridized slides using a CCD-based imaging system (Arrayworx eAuto, Applied Precision) and analyzed them with SoftWoRx Tracker Spot Analysis software. We averaged the ratios of the triplicate spots and calculated standard deviations. All spots with s.d. > 0.075 or signal-to-noise ratios < 20 were removed from the analysis. Repeat experiments showed that results were highly reproducible (fibroblasts $R = 0.88$; SW48 $R = 0.83$). We carried out subsequent data analysis (averaging, ranking, autocorrelation and correlation analysis) in Excel (Microsoft) and S-Plus (Incyte). All microarray data are available for download at our project website.

Calculation of genomic parameters for individual BAC probes. We obtained mapped positions for genes, LINE and Alu elements, and GC percentage (per 20 kb) from the respective tracks on the University of California Santa Cruz Genome Browser version 95 (April 2003 assembly). We used the values and positional information for each category to calculate the gene count, LINE and

Alu densities, and average GC percentage for each of the 32,433 BACs present on the SMRT array. Correlations between gene count and methylation level used averages from 15 probes to account for variable transcript lengths.

CpG island microarray hybridization and analysis. For CpG island array hybridization, we labeled 2 μg of sonicated input DNA with Cy5-dCTP and the product of one MeDIP assay with Cy3-dCTP by random priming using the Bioprime labeling kit (Invitrogen), 120 mM of each dATP, dGTP, dTTP, 60 mM dCTP and 60 mM Cy5-dCTP or Cy3-dCTP. We hybridized the labeled material to the human CpG array 12k from the University Health Network Microarray Centre. This array consists of 12,192 clones derived from a published CpG island library[17]. We carried out hybridization in accordance with the instructions from the University Health Network Microarray Centre. We purified Cy5- and Cy3-labeled probes with the QIAquick PCR purification kit (Qiagen) and mixed them with yeast tRNA (30 μg), salmon sperm DNA (50 μg) and human Cot-1 DNA (25 μg) in DIG Easy Hyb solution (Roche). After denaturation (2 min at 65 °C) and a Cot-1 preannealing step (30 min at 37 °C), we hybridized the slide at 37 °C for 24 h. We washed the array three times in 1× saline sodium citrate and 0.1% SDS at 50 °C and two times in 0.1× saline sodium citrate at room temperature. We scanned the arrays with an Axon 4100B scanner (Axon) and analyzed them using the GenePix Pro 5.0 (Axon) software package and Excel (Microsoft). Features with poor signal-to-noise ratios or saturated pixels were excluded from further analysis. We calculated the ratio between Cy3 and Cy5 signals for all high-quality features and ratio-normalized and $\log_2$-transformed the Cy3 and Cy5 channels using GenePix standard settings. Values are averages of two independent repeats ($R = 0.76$ between WI38 repeats, $R = 0.79$ between SW48 repeats, $R = 0.79$ between colon mucosa repeats). The resulting data sets are accessible from our project website and from the National Center for Biotechnology Information GEO database. We used the following criteria to select hypermethylated clones in SW48 cancer cells relative to fibroblasts or normal colon mucosa: $\log_2$ ratio in SW48 cells > 0.6 and residual $\log_2$ ratio > 0.75. We identified selected clones using the BLAT algorithm on the human genome. Clones with low-quality sequence reads or multiple BLAT hits, as well as those which did not map to CpG islands, were excluded from further analysis.

Bisulfite genomic sequencing. We prepared 200 ng of genomic DNA from SW48 cells, normal colon mucosa and WI38 fibroblasts and embedded it in 25 μl of melted 2% LMP agarose to form beads. We carried out denaturation, treatment with sodium bisulfite, PCR amplification and cloning as previously described[50]. Primer sequences are given in **Supplementary Table 1** online.

***Hpa*II digest.** We digested 2 μg of genomic DNA either with *Xba*I and *Hpa*II or with *Xba*I alone. We carried out PCR reactions on 25 ng of digested DNA using primers spanning fragments that contain several *Hpa*II restriction sites. Primer sequences are given in **Supplementary Table 1** online.

URLs. Our project website is http://www.fmi.ch/members/dirk.schubeler/supplemental.htm. The University Health Network Microarray Centre is available at http://www.microarray.ca/.

GEO accession numbers. CpG island array, GSE2653; SMRT array, GSE2664.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank members of the laboratory of D.S. and W.L.L., C. Alvarez, C. MacAuly and U. Platzbecker for advice; C. Wirbelauer for technical assistance; P. Svoboda for advice on bisulfite genomic sequencing; A. Peters, M. Groudine, M. Lorincz and C. Brown for comments on the manuscript; T. Forné for sharing genomic DNA from hybrid mice; S. Der for access to CpG island sequence reads; B. van Steensel for help in gene annotation; and M. Rebhan for assistance in data analysis. This work was supported by funds from the Novartis Research foundation to D.S.; the Canadian Institute for Health Research, National Institute of Dental Cranial Research, and Genome Canada/British Columbia to W.L.L.; and National Sciences and Engineering Research Council of Canada and Michael Smith Foundation for Health Research Scholarships to J.J.D.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 8 April; accepted 18 May 2005
Published online at http://www.nature.com/naturegenetics/

1. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
2. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33** Suppl., 245–254 (2003).
3. Li, E., Bestor, T.H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
4. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2**, 21–32 (2001).
5. Jones, P.A. & Baylin, S.B. The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3**, 415–428 (2002).
6. Feinberg, A.P. & Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**, 89–92 (1983).
7. Fazzari, M.J. & Greally, J.M. Epigenomics: beyond CpG islands. *Nat. Rev. Genet.* **5**, 446–455 (2004).
8. Rakyan, V.K. *et al.* DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol.* **2**, e405 (2004).
9. Weber, M. *et al.* Extensive tissue-specific variation of allelic methylation in the Igf2 gene during mouse fetal development: relation to expression and imprinting. *Mech. Dev.* **101**, 133–141 (2001).
10. Tremblay, K.D., Duran, K.L. & Bartolomei, M.S. A 5′ 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol. Cell. Biol.* **17**, 4322–4329 (1997).
11. Ishkanian, A.S. *et al.* A tiling resolution DNA microarray with complete coverage of the human genome. *Nat. Genet.* **36**, 299–303 (2004).
12. Migeon, B.R. X-chromosome inactivation: molecular mechanisms and genetic consequences. *Trends Genet.* **10**, 230–235 (1994).
13. Bernardino, J., Lombard, M., Niveleau, A. & Dutrillaux, B. Common methylation characteristics of sex chromosomes in somatic and germ cells from mouse, lemur and human. *Chromosome Res.* **8**, 513–525 (2000).
14. Viegas-Pequignot, E., Dutrillaux, B. & Thomas, G. Inactive X chromosome has the highest concentration of unmethylated Hha I sites. *Proc. Natl. Acad. Sci. USA* **85**, 7657–7660 (1988).
15. Giacalone, J., Friedes, J. & Francke, U. A novel GC-rich human macrosatellite VNTR in Xq24 is differentially methylated on active and inactive X chromosomes. *Nat. Genet.* **1**, 137–143 (1992).
16. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
17. Cross, S.H., Charlton, J.A., Nan, X. & Bird, A.P. Purification of CpG islands using a methylated DNA binding column. *Nat. Genet.* **6**, 236–244 (1994).
18. Groot, G.S. & Kroon, A.M. Mitochondrial DNA from various organisms does not contain internally methylated cytosine in -CCGG- sequences. *Biochim. Biophys. Acta* **564**, 355–357 (1979).
19. Yan, P.S. *et al.* Hypermethylation of ribosomal DNA in human breast carcinoma. *Br. J. Cancer* **82**, 514–517 (2000).
20. Oakes, C.C., Smiraglia, D.J., Plass, C., Trasler, J.M. & Robaire, B. Aging results in hypermethylation of ribosomal DNA in sperm and liver of male rats. *Proc. Natl. Acad. Sci. USA* **100**, 1775–1780 (2003).
21. Salem, C.E. *et al.* PAX6 methylation and ectopic expression in human tumor cells. *Int. J. Cancer* **87**, 179–185 (2000).
22. Yan, P.S. *et al.* CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer. *Clin. Cancer Res.* **6**, 1432–1438 (2000).
23. Louro, R. *et al.* RASL11A, member of a novel small monomeric GTPase gene family, is down-regulated in prostate tumors. *Biochem. Biophys. Res. Commun.* **316**, 618–627 (2004).
24. Watson, J.E. *et al.* Integration of high-resolution array comparative genomic hybridization analysis of chromosome 16q with expression array data refines common regions of loss at 16q23-qter and identifies underlying candidate tumor suppressor genes in prostate cancer. *Oncogene* **23**, 3487–3494 (2004).
25. Derynck, R., Akhurst, R.J. & Balmain, A. TGF-beta signaling in tumor suppression and cancer progression. *Nat. Genet.* **29**, 117–129 (2001).
26. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. USA* **96**, 8681–8686 (1999).
27. Suter, C.M., Martin, D.I. & Ward, R.L. Germline epimutation of MLH1 in individuals with multiple cancers. *Nat. Genet.* **36**, 497–501 (2004).
28. Adorjan, P. *et al.* Tumour class prediction and discovery by microarray-based DNA methylation analysis. *Nucleic Acids Res.* **30**, e21 (2002).
29. Rougier, N. *et al.* Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108–2113 (1998).
30. Rabinowicz, P.D. *et al.* Genes and transposons are differentially methylated in plants, but not in mammals. *Genome Res.* **13**, 2658–2664 (2003).
31. Bird, A.P. & Wolffe, A.P. Methylation-induced repression–belts, braces, and chromatin. *Cell* **99**, 451–454 (1999).
32. Rountree, M.R. & Selker, E.U. DNA methylation inhibits elongation but not initiation of transcription in Neurospora crassa. *Genes Dev.* **11**, 2383–2395 (1997).
33. Lorincz, M.C., Dickerson, D.R., Schmitt, M. & Groudine, M. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat. Struct. Mol. Biol.* **11**, 1068–1075 (2004).

34. Bird, A.P. Gene number, noise reduction and biological complexity. *Trends Genet.* **11**, 94–100 (1995).
35. Walsh, C.P., Chaillet, J.R. & Bestor, T.H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117 (1998).
36. Woodfine, K. *et al.* Replication timing of the human genome. *Hum. Mol. Genet.* **13**, 191–202 (2004).
37. White, E.J. *et al.* DNA replication-timing analysis of human chromosome 22 at high resolution and different developmental states. *Proc. Natl. Acad. Sci. USA* **101**, 17771–17776 (2004).
38. Gilbert, N. *et al.* Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* **118**, 555–566 (2004).
39. Chadwick, B.P. & Willard, H.F. Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proc. Natl. Acad. Sci. USA* **101**, 17450–17455 (2004).
40. Kohlmaier, A. *et al.* A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. *PLoS Biol.* **2**, E171 (2004).
41. Gaudet, F. *et al.* Induction of tumors in mice by genomic hypomethylation. *Science* **300**, 489–492 (2003).
42. Xu, G.L. *et al.* Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187–191 (1999).
43. Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L. & Jaenisch, R. DNA hypomethylation leads to elevated mutation rates. *Nature* **395**, 89–93 (1998).
44. Costello, J.F. *et al.* Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.* **24**, 132–138 (2000).
45. Lee, T.I. *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
46. Cam, H. *et al.* A common set of gene regulatory networks links metabolism and growth inhibition. *Mol. Cell* **16**, 399–411 (2004).
47. Pollack, J.R. *et al.* Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23**, 41–46 (1999).
48. Schubeler, D. *et al.* The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18**, 1263–1271 (2004).
49. Schubeler, D. *et al.* Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. *Nat. Genet.* **32**, 438–442 (2002).
50. Svoboda, P., Stein, P., Filipowicz, W. & Schultz, R.M. Lack of homologous sequence-specific DNA methylation in response to stable dsRNA expression in mouse oocytes. *Nucleic Acids Res.* **32**, 3601–3606 (2004).