PERSPECTIVE

nature genetics

Evidence for an instructive mechanism of *de novo* methylation in cancer cells

Ilana Keshet, Yeshayahu Schlesinger, Shlomit Farkash, Eyal Rand, Merav Hecht, Eran Segal, Eli Pikarski, Richard A Young, Alain Niveleau, Howard Cedar & Itamar Simon

DNA methylation has a role in the regulation of gene expression during normal mammalian development but can also mediate epigenetic silencing of CpG island genes in cancer and other diseases. Many individual genes (including tumor suppressors) have been shown to undergo *de novo* methylation in specific tumor types, but the biological logic inherent in this process is not understood. To decipher this mechanism, we have adopted a new approach for detecting CpG island DNA methylation that can be used together with microarray technology. Genome-wide analysis by this technique demonstrated that tumor-specific methylated genes belong to distinct functional categories, have common sequence motifs in their promoters and are found in clusters on chromosomes. In addition, many are already repressed in normal cells. These results are consistent with the hypothesis that cancer-related *de novo* methylation may come about through an instructive mechanism.

In the mammalian genome, DNA methylation occurs in a stepwise manner during normal development. Initially, almost all methyl groups inherited from the gametes are erased in the morula[1], but the basic methylation pattern is then re-established in each individual at about the time of implantation[2]. Although the entire genome is subject to a general wave of *de novo* methylation at this stage, CpG islands are specifically recognized and protected from this modification by virtue of common *cis*-acting sequences[3,4]. Once the bimodal pattern is set up by this global mechanism, select CpG islands may still undergo *de novo* methylation at a later stage, but this is then carried out by a targeting mechanism through which *de novo* methylases are recruited to specifically recognized regions of the genome. It is this type of mechanism, for example, that is responsible for the *de novo* methylation of CpG islands on the inactive X chromosome in female cells[5] and the *Pouf5f1* (also known as *Oct-3/4*) gene promoter when it undergoes silencing in the post-implantation embryo[6]. Once these dynamic changes are established, the resulting methylation pattern is then preserved through subsequent cell divisions by virtue of maintenance methylation, even though the original stage-specific factors involved in its establishment may no longer be present.

Ilana Keshet, Yeshayahu Schlesinger, Eyal Rand, Merav Hecht and Howard Cedar are at the Department of Cellular Biochemistry and Human Genetics; Shlomit Farkash and Itamar Simon are at the Department of Molecular Biology; and Eli Pikarski is at the Department of Pathology of Hebrew University, Jerusalem, Israel. Eran Segal is at the Center for Studies in Physics and Biology, Rockefeller University, New York, New York, USA. Richard A. Young is at the Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts, USA. Alain Niveleau is at the Laboratoire de Virologie Structurale et Moleculaire, Universite Fourier de Grenoble, Avenue Gresivandan, La Tronche, France.
e-mail: cedar@md.huji.ac.il

Published online 24 January 2006 ; doi:10.1038/ng1719

DNA methylation in cancer

A large body of evidence indicates that many CpG islands can undergo abnormal *de novo* methylation in cancer[7–9]. Initially, this was shown to be part of a silencing mechanism for tumor suppressor–like genes[10], but subsequent experiments demonstrated that this phenomenon is probably more widespread, clearly encompassing a large variety of different gene types[11–13] that are not necessarily involved in cell growth or tumorigenesis. This aberrant methylation generates a discrete cell type– and gene-specific pattern[14] that seems to be set up relatively early in tumor development[15]. Indeed, many of the affected genes have been demonstrated to actually undergo *de novo* methylation in normal tissues as a function of aging[16,17].

It is commonly accepted that singular genetic changes observed in cancer cells come about in a random manner but are then selected because they lead to an enhanced growth phenotype. In keeping with this standard model, it has been suggested that epigenetic alterations such as *de novo* methylation are also generated in a nonspecific manner, perhaps through a general leakiness in the basic repression machinery of the cell, with the final pattern being determined cumulatively by selection. Alternatively, specific genes or regions of the genome may actually be targeted for repression through a biologically programmed *trans*-acting instructive pathway[18]. To a large extent, this key question might be resolved if one had a better picture of the full complement of genes that actually undergo *de novo* methylation.

Genome-wide detection of *de novo* methylation

Although there are a number of excellent techniques available for mapping DNA methylation on individual gene sequences, there are very few methods that can be used to detect DNA methylation genome-wide. The restriction landmark genomic scanning (RLGS) system, for example, has proven useful for picking up methylation differences in a variety of tumors[19], but this method is limited to detecting methyl groups at a small number of unique restriction sites, requires specialized equipment and is not readily amenable to gene identification[20]. A very promising alternate

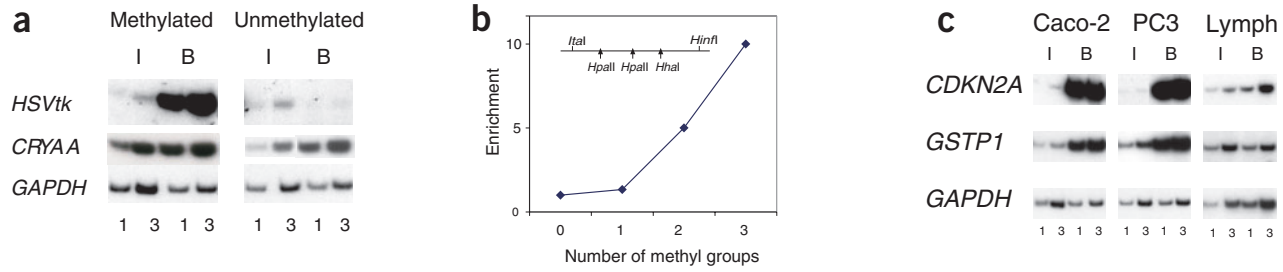

Figure 1 Detection of DNA methylation by mDIP. (**a**) The *HSVtk* plasmid[39] was methylated *in vitro* with *Sss*I methylase (New England Biolabs). Methylated or unmethylated plasmid (1 μg) was added to DNA from human lymphoblasts, sonicated and subjected to mDIP. Semiquantitative PCR was carried out on input (I) and bound (B) DNA using 1 and 3 μl samples. Phosphorimager analysis indicated that the methylated *HSVtk* sequence was enriched (B/I) more than 100-fold when normalized to the human *GAPDH* promoter. In contrast, an unmethylated *HSVtk* sequence was not enriched (0.7-fold) compared with a positive methylated control (*CRYAA*) that was enriched five- to eightfold. (**b**) A 100 bp fragment containing either one, two or three methyl groups was subject to mDIP analysis (**Supplementary Methods**). These results indicate that the sensitivity of this technique for detecting DNA methylation is about 2–3%. (**c**) DNA samples from colon (Caco-2) or prostate (PC3) cancer cell lines and normal lymphoblasts were sonicated, subjected to mDIP and assayed for various gene sequences. Phosphorimager analysis indicates that *CDKN2A* is enriched more than 100-fold and *GSTP1* 30-fold in both Caco-2 and PC3 as compared with *GAPDH* in each cell type. In contrast, these genes were enriched less than twofold in lymphoblast cells.

approach is to use immunoprecipitation to identify methylated sequences in the genome[21,22].

On the basis of this idea, we have used a methyl-DNA immunoprecipitation (mDIP) assay that uses antibodies specific for 5-methyl–cytosine residues. To demonstrate the feasibility of this method, we methylated an *HSVtk* plasmid *in vitro* at every CpG residue using the *Sss*I methylase, added an excess of human cellular DNA and then carried out immunoprecipitation using an antibody specific for 5-methylcytosine. We amplified input and bound DNA by sequence-specific PCR and determined the amount of enrichment after normalizing for the unmethylated *GAPDH* CpG island promoter region. PCR analysis showed that the methylated *HSVtk* sequence was indeed highly enriched, whereas the unmethylated *HSVtk* control appeared depleted (**Fig. 1a**). To obtain a better estimation for the resolution of this method, we devised an *in vitro* plasmid system for measuring the degree of enrichment as a function of the number of methyl groups on the DNA. This analysis clearly showed that although one methyl group is insufficient, the presence of two or three methyl moieties per molecule is easily detected by this technique (**Fig. 1b**).

In order to determine whether this assay could be used to detect abnormal *de novo* methylation in cancer cells, we carried out mDIP on DNA from several tumor cell lines and used PCR to detect methylation at cancer-related genes (**Fig. 1c**). Notably, we found that both the *CDKN2A* (also known as *p16*) and *GSTP1* promoters were highly enriched for methyl groups in both colon carcinoma (Caco-2) and prostate cancer (PC3) cells, as has been demonstrated previously[23,24]. We did not observe any enrichment using DNA from a normal lymphoblast cell line[23,25].

We next attempted to use mDIP to detect DNA methylation on a genome-wide basis. As a first step, we carried out mDIP on normal lymphoblasts, labeled the input DNA with a cyan-3 dye, labeled the bound DNA with cyan-5 and then hybridized the mixture to a microarray chip carrying PCR products from a collection of ~13,000 human gene promoters[26]. Under these conditions, methylated sequences are enriched in cyan-5 and therefore show a red signal, whereas unmethylated DNA is predominantly green (**Supplementary Fig. 1** online). A careful analysis of these results indicated that the antibody used in this experiment recognizes methylation only in regions with a threshold CpG density of about 2–3% (**Supplementary Fig. 1**). Although this might not be sensitive enough for mapping methylation in most genomic regions where the CpG content is generally less than 2%, this method is ideally suited for detecting abnormally methylated CpG islands.

De novo methylation in cancer cells

Using this assay, we then carried out genome-wide mDIP on DNA from the colon cancer cell line Caco-2. In order to maximize our ability to detect aberrant methylation, we compared the bound fraction from Caco-2 with the bound fraction from normal lymphoblast cells, in which almost all promoter CpG islands are unmethylated. In this assay, we found a set of 135 gene promoters that were specifically methylated in the cancer cell line ($P < 0.001$); 127 (94%) of these actually contained CpG islands (**Fig. 2a**). We obtained similar results in repeat experiments, or when the bound fraction from Caco-2 was compared with the bound fraction from normal colon (**Fig. 2b**). This is consistent with the idea that almost all CpG islands are constitutively unmethylated in every cell type. Indeed, a microarray comparison between normal colon and lym-

Table 1 Functional analysis of methylated genes in cancer cells

Enriched categories	Caco-2 (105)	PC3 (88)	Tumors (137)
Cell adhesion (213)	**12** (0.004)	**6** (0.003)	**20** (5.5×10^{-6})
Homophilic cell adhesion (56)	**9** (3.8×10^{-6})	**6** (0.0006)	**13** (5.7×10^{-9})
Cell-cell signaling (204)	**14** (0.0001)	**11** (0.001)	**20** (1.4×10^{-6})
Synaptic transmission (102)	**8** (0.002)	**8** (0.0007)	**14** (1.5×10^{-6})
Signal transduction (945)	**36** (0.0007)	**28** (0.009)	**45** (0.0005)
G protein–coupled receptor protein signaling pathway (182)	**15** (1.4×10^{-5})	**15** (1.5×10^{-6})	**24** (4.4×10^{-10})
Second messenger–mediated signaling (65)	**4**	**5** (0.007)	**7** (0.002)
Ion transport (247)	**14** (0.001)	**9**	**17** (0.0008)

Genes enriched for DNA methylation ($P < 0.001$) in Caco-2, PC3 or colon carcinoma were organized into functional groups according to the Gene Ontology (GO) annotation. The number of genes in each category is shown, as well as the hypergeometric *P* value (in parentheses). These were compared with the number of genes (in parentheses) in each category out of a total of 4,494 genes with known function (**Supplementary Table 3** online contains an expanded version of this table). The data from colon tumors was derived from all 170 genes deemed methylated in any of the six samples tested (only 137 of them are in the GO database).

phoblast cells demonstrated only a few differentially methylated CpG island sequences (**Supplementary Methods** online). mDIP analysis of the prostate cancer cell line, PC-3, also demonstrated many *de novo* methylated genes, and although some of these were detected in Caco-2, others seemed to be unique to the prostate cancer phenotype (**Fig. 2**).

In order to validate these results at the molecular level, we randomly chose ten of the presumed methylated promoters and subjected them to whole-population bisulfite analysis (**Supplementary Fig. 2** online). Notably, almost all (>90%) of the 164 CpG sites included in this screen were highly methylated (that is, methylated on both alleles) in Caco-2 DNA and either unmethylated or only slightly methylated (10–30%) in lymphoblast cells and normal colon (**Supplementary Methods**). Previous studies have shown that many of the genes detected by this method are methylated in tumor cells and are inducible by 5azaC treatment (**Supplementary Table 1** online). When taken together, these results indicate that mDIP microarray represents a good technique for detecting genome-wide *de novo* methylation of CpG islands in cancer cells.

In order to confirm that the results in cell lines were representative of tumor tissues themselves, we carried out an mDIP microarray analysis comparing bound DNA from six different colon tumor samples to bound DNA from normal lymphocytes or colon (**Fig. 2c** and **Supplementary Table 2** online). Here, too, a considerable number of genes were abnormally methylated ($P < 0.001$). Although some of these were unique to the tumor DNA, the majority of genes were similar to those found modified in Caco-2 cells (**Fig. 2d**). As already shown in previous studies, the number of methyl groups observed on any particular gene sequence in tumors was generally less than that observed in cell lines[27] partly, but not only, because these samples may have been contaminated with unmethylated DNA from normal surrounding cell types. This general idea was confirmed when we highlighted the Caco-2 methylated genes on the scatter plot for each tumor (**Fig. 2c**). Almost all of these sequences (85%) were skewed in the direction of methyl enrichment.

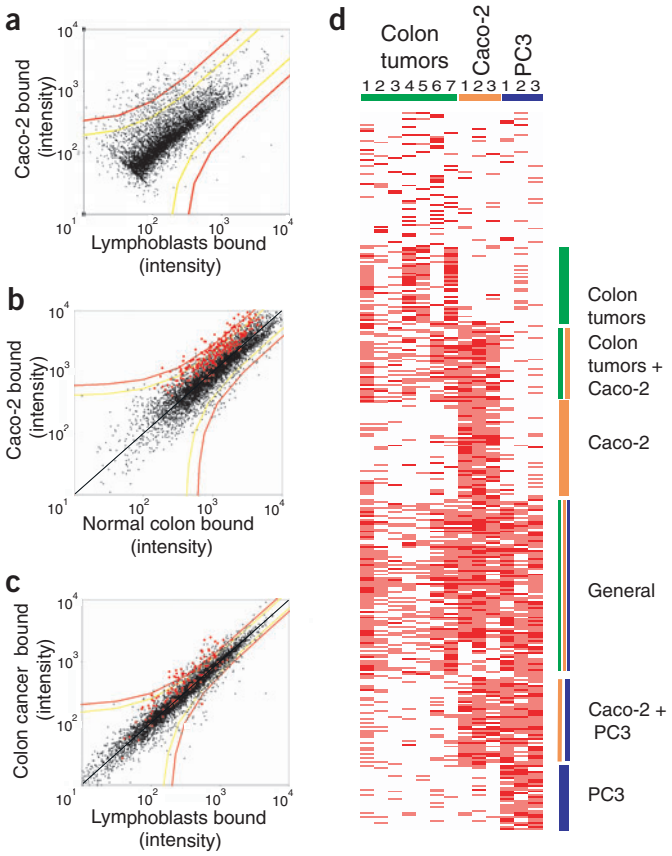

Figure 2 mDIP microarray analysis. The intensity of cyan-3 and cyan-5 for each spot is displayed as a scatter plot. Boundary lines indicate confidence limits of $P < 0.01$ (yellow) or $P < 0.001$ (red). The full data set for these experiments is presented in **Supplementary Table 4** online and can be downloaded from the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress; accession number E-MEXP-520. (**a**) Bound Caco-2 DNA versus bound lymphoblast DNA. Differentially methylated genes ($P < 0.001$) appear above the upper red line. (**b**) Bound Caco-2 DNA versus bound DNA from normal colon. Red dots represent 127 genes originally found hypermethylated in Caco-2 DNA but not in DNA from normal lymphoblasts. (**c**) Bound carcinoma DNA versus bound DNA from normal lymphocytes (black dots). Red dots represent 127 genes found hypermethylated in Caco-2 ($P < 0.001$). (**d**) Methylation state of 367 genes detected as being methylated ($P < 0.001$) at least once in assays of Caco-2, PC3 or colon tumor samples is shown for each mDIP microarray experiment. We analyzed six individual colon tumors, comparing three polyps against lymphocyte DNA (1–3) and one against normal colon (4) and three carcinomas against lymphocyte DNA (5–7). We compared Caco-2 with lymphoblast DNA (1–2) or normal colon (3) and compared PC3 with lymphoblast DNA in repeat experiments (1–3). Red indicates strong methylation ($P < 0.001$), whereas pink indicates moderate methylation ($0.001 < P < 0.1$). The genes have been clustered according to cell type specificity. Genes specifically methylated in colon tumors (green), Caco-2 (orange) or PC3 (blue) are shown. Many genes are generally methylated in all cancer cells.

Functional analysis

The microarray used in this study contains approximately 10,000 verified promoter elements and thus represents only a portion of the genome. Nonetheless, this promoter set is extensive enough to serve as the basis for bioinformatics analysis. Using three different approaches, we asked whether methylated gene promoters might have common characteristics. As a first step, we divided the methylated genes from Caco-2 cells according to biological function (**Table 1** and **Supplementary Table 3** online). This demonstrated a highly significant enrichment of methylated promoters of genes in four functional categories: (i) cell adhesion (ii) signal transduction (iii) ion transport and (iv) cell-cell signaling. Furthermore, this seems to be true of *de novo* methylation in general, as methylated genes from PC-3, as well as the set of CpG island sequences deemed methylated in colon tumor tissue, were also enriched in these same functional categories (**Table 1**). Notably, this analysis did not detect preferential enrichment for genes involved in cell cycle control or apoptosis, as might have been expected if growth selection were to play a major role in molding the *de novo* methylation pattern.

Motif analysis

As a second approach, we used an algorithm[28] to search for common sequence elements. Here, too, we identified several short sequence motifs that were significantly enriched in methylated promoters as compared with the total population of CpG island promoters, both in Caco-2 and PC3 cells (**Fig. 3a**). Although these motifs have not been characterized as recognition sites for protein binding, their overrepresentation in methylated promoters suggests that they may have a role in the targeting process. As all of these enriched sequences seem to be abundant in the general promoter population, however, it is clear that any effect on targeting must be in collaboration with other factors.

Chromosomal distribution

In another test for common features, we analyzed the chromosomal distribution of the methyl-targeted genes (**Fig. 3b**). About 20% of the enriched sequences in Caco-2 are located in distinct clusters containing two or more adjacent genes. In one instance, four methylated gene sequences were found to be located within a single 1-Mb cluster on

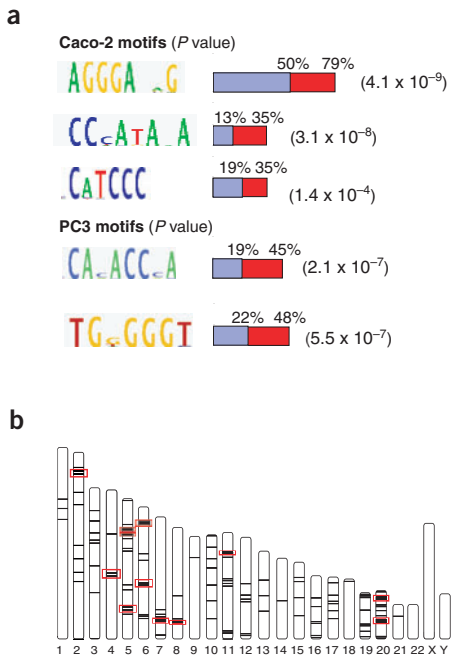

Figure 3 Genomic analysis of methylated genes in cancer cells. (**a**) Promoter sequences that are statistically enriched in hypermethylated promoter regions from Caco-2 and PC3. The relative abundance of each nucleotide is depicted by the height of each letter. The bar graph shows the percentage of the methylated (red) or total promoter (blue) population with each motif, together with its hypergeometric *P* value. Some of the motifs enriched in Caco-2 methylated genes are also enriched in the collection of 191 tumor methylated genes. As a control we ran the motif program on a random set of CpG island promoters and did not detect any statistically significant motifs. (**b**) Chromosome map showing the positions of all 127 genes that are hypermethylated in Caco-2 ($P < 0.001$). Regions that contain clustered adjacent genes (two or more) are outlined in red. The number of clusters is highly significant ($P = 0.0003$). This clustering is observed despite the fact that the mDIP input DNA was sonicated to an average size of 700 bp, much smaller than the distance between adjacent genes. Clustering ($P < 0.00006$) was also observed for the set of all methylated genes (367).

chromosome 5. These results are highly significant in light of the fact that the probability of finding this type of clustering in a random sample of 127 genes is extremely low ($P = 0.0003$). A similar clustered distribution ($P = 0.00006$) was also observed for the full set of 367 genes found to be methylated (**Fig. 2d**) in our assay. This type of nonrandom topographic organization raises the possibility that *de novo* methylation in cancer might be targeted to select genomic domains through the recognition of *cis*-acting elements or epigenetic markers that act in a regional manner.

Gene expression analysis

To obtain a better picture of the genes that become methylated in colon carcinoma cells, we investigated expression levels of these genes in normal and tumor tissues. Data derived from the gene atlas project[29] indicated that only 15 of the genes that were *de novo* methylated in Caco-2 cells are defined as active (**Supplementary Methods**) in normal colon (appendix). These particular genes were found to be highly repressed in a collection of colon carcinoma samples (**Fig. 4a**). In addition, we found that this was true even for individual precancerous adenomas, when they were compared with normal tissue[30] (**Fig. 4b**). This latter observation, together with the fact that methylation is already detected in polyps (**Fig. 2d**), strongly suggests that the full pattern of *de novo* modification may actually be generated early in tumorigenesis, as has been previously noted for select genes[14].

In contrast to these active sequences, most (91/106) of the genes are already expressed at relatively low levels in normal colon and do not undergo further repression in cancer. We observed a similar expression pattern in tumor cells and normal cells for 170 genes found methylated in colon tumor samples (**Fig. 4c**) and for the 104 methylated genes in PC3 (**Fig. 4d**). Thus, it seems that *de novo* methylation is not necessarily responsible for the repression of every target gene. In terms of understanding the logic of this process, these observations clearly imply that a considerable number of *de novo* events may not be subject to growth selection.

Lessons from bioinformatics analysis

Most studies on DNA methylation in cancer have focused on individual cases of tumor suppressors or other gene sequences involved in the control of cell growth or apoptosis. Although a number of studies have attempted to detect additional gene targets (see refs. 19, 21), in general, their methodologies have not been sensitive enough to identify a representative population of defined gene sequences large enough for bioinformatics analysis. By using mDIP technology in combination with a relatively large promoter microarray, it has been possible for the first time to identify a large number of target genes in a completely unbiased manner, allowing examination of the logic of *de novo* methylation.

Our analysis shows that many methylated genes are located within defined genomic clusters, are associated with common sequence motifs and belong to specific functional categories. These findings strongly suggest that this modification process does not occur in a random manner; this is further supported by the observation that methylation takes place equally

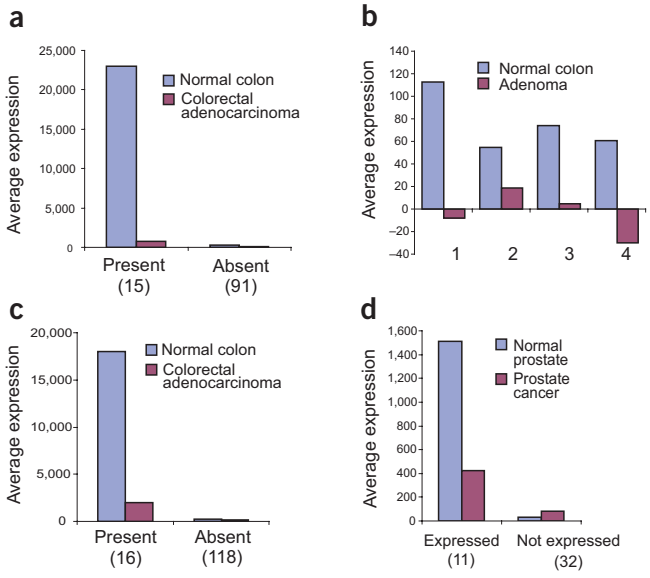

Figure 4 Effect of methylation on gene expression. (**a**) Average level of gene expression (using the gene atlas V2 dataset) in normal colon tissue (appendix) as compared with carcinoma for the Caco-2 methylated genes that are categorized as either present or absent in normal appendix. The number of genes in each category is shown in parentheses. (**b**) Matched comparison between four samples of colon adenoma and normal colon from the same patients for the average expression levels of the Caco-2 methylated genes that are expressed (activity >500) in normal colon (appendix)[29]. (**c**) Average level of gene expression (as in **a**) for genes methylated in colon tumor samples (**Fig. 2d**). (**d**) Average level of gene expression in normal prostate versus the prostate cancer LNCaP cell line (downloaded from the NCBI GEO database; accession numbers GSM12743 and GSM11873, respectively) for the PC3-methylated genes that are expressed (activity >500) or not expressed in normal prostate cells[29].

on both alleles of almost all targeted genes[12] (**Supplementary Fig. 2**). In addition, our analysis indicates that many methylated genes are not necessarily involved in the control of cell proliferation and may actually be inactive in normal tissue samples before tumor formation, making it unlikely that these methylation patterns are dictated solely by growth selection.

Our data do not necessarily imply that all cancer-related methylation is preprogrammed and independent of selection. For many tumor suppressor–like genes, for example, *de novo* methylation clearly has been shown to serve as an epigenetic silencing mechanism[10]. Furthermore, in almost all tumors that carry a germline mutation of these genes, second-step DNA promoter methylation always occurs exclusively on the wild-type allele[31], strongly suggesting that this particular form of modification may come about as part of the process of growth selection. This mechanism is clearly unlike that involved in the generally biallelic methylation of target genes detected using nonbiased microarray (**Fig. 2**) or RLGS[12] analysis.

Instructive mechanism

On the basis of this bioinformatics analysis, we would like to raise the hypothesis that most *de novo* methylation in cancer may take place in an instructive manner through interactions between *cis*-acting sequences on the DNA and *trans*-acting protein complexes capable of recruiting DNA methyltransferases. A good example of this mechanism has been observed in acute promyelocytic leukemia: the PML-RAR fusion protein can induce gene hypermethylation and gene silencing at specific target promoters[32]. Overexpression of DNA methyltransferase is itself capable of bringing about the *de novo* methylation of selective CpG islands[33], a high percentage of which may be methylated in cancer (see preliminary results in **Supplementary Methods**), suggesting that specific genes may be marked as preferred substrates for modification.

One potential candidate for a factor that may be capable of directing targeted methylation in cancer is the polycomb repression protein EZH2, which is overexpressed in a variety of different human tumors[34,35], is involved in the targeted *de novo* methylation that takes place during X chromosome inactivation in normal embryos[36,37] and has recently been shown to be able to recruit DNA methyltransferases[38]. Although additional studies will be required to further define these *cis*- and *trans*-acting factors, now it may be possible to use the functional (**Table 1**) and structural (**Fig. 3a**,**b**) information obtained from genome-wide methylation analysis to decipher the biological pathways and molecular mechanisms involved in programming *de novo* methylation.

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

ACKNOWLEDGMENTS

Colon biopsies were provided by J. Zimmerman. We thank R. Rubinstein and M. Kott for help in the computer and microarray analyses. This research was supported by grants from the Rosetrees Trust, the US National Institutes of Health, the Israel Cancer Research Fund, the Belfer Foundation, the Conquer Fragile X Foundation, the Prostate Cancer Foundation, Fondation pour la Recherche Medicale-Ligue Nationale Francaise Contre le Cancer, Association for International Cancer Research (AICR) and the Bikura program of the Israel Science Foundation.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382 (1987).
2. Kafri, T. *et al.* Developmental pattern of gene-specific DNA methylation in the mouse embryo and germline. *Genes Dev.* **6**, 705–714 (1992).
3. Brandeis, M. *et al.* Sp1 elements protect a CpG island from de novo methylation. *Nature* **371**, 435–438 (1994).
4. Macleod, D., Charlton, J., Mullins, J. & Bird, A.P. Sp1 sites in the mouse Aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* **8**, 2282–2292 (1994).
5. Lock, L.F., Takagi, N. & Martin, G.R. Methylation of the HPRT gene on the inactive X occurs after chromosome inactivation. *Cell* **48**, 39–46 (1987).
6. Gidekel, S. & Bergman, Y. A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element. *J. Biol. Chem.* **277**, 34521–34530 (2002).
7. Robertson, K.D. & Wolffe, A.P. DNA methylation in health and disease. *Nat. Rev. Genet.* **1**, 11–19 (2000).
8. Baylin, S.B. & Herman, J.G. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.* **16**, 168–174 (2000).
9. Jones, P.A. & Laird, P.W. Cancer epigenetics comes of age. *Nat. Genet.* **21**, 163–167 (1999).
10. Jones, P.A. & Baylin, S.B. The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3**, 415–428 (2002).
11. Costello, J.F. *et al.* Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.* **24**, 132–138 (2000).
12. Zardo, G. *et al.* Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. *Nat. Genet.* **32**, 453–458 (2002).
13. Yu, L. *et al.* Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. *Nat. Genet.* **37**, 265–274 (2005).
14. Issa, J.P. CpG island methylator phenotype in cancer. *Nat. Rev. Cancer* **4**, 988–993 (2004).
15. Chan, A.O. *et al.* CpG island methylation in aberrant crypt foci of the colorectum. *Am. J. Pathol.* **160**, 1823–1830 (2002).
16. Toyota, M. *et al.* Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. *Cancer Res.* **59**, 5438–5442 (1999).
17. Ahuja, N. & Issa, J.P. Aging, methylation and cancer. *Histol. Histopathol.* **15**, 835–842 (2000).
18. Baylin, S. & Bestor, T.H. Altered methylation patterns in cancer cell genomes: cause or consequence? *Cancer Cell* **1**, 299–305 (2002).
19. Rush, L.J. & Plass, C. Restriction landmark genomic scanning for DNA methylation in cancer: past, present, and future applications. *Anal. Biochem.* **307**, 191–201 (2002).
20. Smiraglia, D.J. & Plass, C. The study of aberrant methylation in cancer via restriction landmark genomic scanning. *Oncogene* **21**, 5414–5426 (2002).
21. Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853–862 (2005).
22. Mukhopadhyay, R. *et al.* The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res.* **14**, 1594–1602 (2004).
23. Toyota, M. *et al.* Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.* **59**, 2307–2312 (1999).
24. Stirzaker, C., Song, J.Z., Davidson, B. & Clark, S.J. Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells. *Cancer Res.* **64**, 3871–3877 (2004).
25. Maruyama, R. *et al.* Aberrant promoter methylation profile of prostate cancers and its relationship to clinicopathological features. *Clin. Cancer Res.* **8**, 514–519 (2002).
26. Odom, D.T. *et al.* Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381 (2004).
27. Suzuki, H. *et al.* A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat. Genet.* **31**, 141–149 (2002).
28. Segal, E., Yelensky, R. & Koller, D. Genome-wide discovery of transcriptional modules from DNA sequence and gene expression. *Bioinformatics* **19** (Suppl.), i273–i282 (2003).
29. Su, A.I. *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. USA* **101**, 6062–6067 (2004).
30. Notterman, D.A., Alon, U., Sierk, A.J. & Levine, A.J. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. *Cancer Res.* **61**, 3124–3130 (2001).
31. Esteller, M. *et al.* DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum. Mol. Genet.* **10**, 3001–3007 (2001).
32. Di Croce, L. *et al.* Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295**, 1079–1082 (2002).
33. Feltus, F.A., Lee, E.K., Costello, J.F., Plass, C. & Vertino, P.M. Predicting aberrant CpG island methylation. *Proc. Natl. Acad. Sci. USA* **100**, 12253–12258 (2003).
34. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).
35. Cao, R. & Zhang, Y. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr. Opin. Genet. Dev.* **14**, 155–164 (2004).
36. Plath, K. *et al.* Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300**, 131–135 (2003).
37. Silva, J. *et al.* Establishment of histone H3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev. Cell* **4**, 481–495 (2003).
38. Vire, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* advance online publication 14 December 2005 (doi:10.1038/nature04431).
39. Eden, S., Hashimshony, T., Keshet, I., Thorne, A.W. & Cedar, H. DNA methylation models histone acetylation. *Nature* **394**, 842–843 (1998).