

# Bivalent domains enforce transcriptional memory of DNA methylated genes in cancer cells

Jairo Rodriguez<sup>a,b</sup>, Mar Muñoz<sup>a</sup>, Laura Vives<sup>a</sup>, Costas G. Frangou<sup>c</sup>, Mark Groudine<sup>c,1</sup>, and Miguel A. Peinado<sup>a,b,2</sup>

<sup>a</sup>Institut de Medicina Predictiva i Personalitzada del Càncer, Badalona, 08916 Barcelona, Spain; <sup>b</sup>Institut d'Investigació Biomèdica de Bellvitge, L'Hospitalet, 08907 Barcelona, Spain; and <sup>c</sup>Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA

Contributed by Mark T. Groudine, October 8, 2008 (sent for review August 21, 2008)

**Silencing of multiple cancer-related genes is associated with de novo methylation of linked CpG islands. Additionally, bivalent histone modification profiles characterized by the juxtaposition of active and inactive histone marks have been observed in genes that become hypermethylated in cancer. It is unknown how these ambiguous epigenetic states are maintained and how they interrelate with adjacent genomic regions with different epigenetic landscapes. Here, we present the analysis of a set of neighboring genes, including many frequently silenced in colon cancer cells, in a chromosomal region at 5q35.2 spanning 1.25 Mb. Promoter DNA methylation occurs only at genes maintained at a low transcriptional state and is characterized by the presence of bivalent histone marks, namely trimethylation of lysines 4 and 27 in histone 3. Chemically induced hyperacetylation and DNA demethylation lead to up-regulation of silenced genes in this locus yet do not resolve bivalent domains into a domain-wide active chromatin conformation. In contrast, active genes in the region become down-regulated after drug treatment, accompanied by a partial loss of chromatin domain boundaries and spreading of the inactive histone mark trimethylated lysine 27 in histone 3. Our results demonstrate that bivalent domains mark the promoters of genes that will become DNA methylated in adult tumor cells to enforce transcriptional silence. These bivalent domains not only remain upon drug induced gene reactivation, but also spread over adjacent CpG islands. These results may have important implications in understanding and managing epigenetic therapies of cancer.**

colorectal cancer | DNA methylation | epigenetic memory | transcriptional silencing

It is now clear that epigenetic events, in cooperation with genetic events, are involved in every step of tumorigenesis and play a critical role in the disruption of key cellular pathways deregulated in human cancers (1, 2).

De novo methylation of CpG islands is associated with the transcriptional silencing of many cancer-related genes (1, 2). The promoter regions of silenced genes, including those with promoter DNA methylation, contain specific histone modifications, which are a signature of transcriptional inactivation (3). Additionally, the DNA methylation mark itself can be read by specific proteins that can alter chromatin structure (4). Thus, a cross-talk exists between DNA methylation and histone modifications to orchestrate transcriptional silencing.

A growing body of evidence suggests that the two main cell memory systems implicated in the maintenance of a stem cell state, Trithorax (Trx) and Polycomb Group (PcG) proteins, may be involved in tumor-associated aberrant gene silencing and promoter DNA methylation (5, 6). In this context, the active mark, methylated lysine 4, together with the silent mark, methylated lysine 27, have been found to coexist over the promoter regions of DNA methylated genes in human cancer cells (7). This epigenetic landscape is similar to the bivalent domains characterized by the concurrence of trimethylated lysine 4 (H3K4me3) and trimethylated lysine 27 marks (H3K27me3), which have been described for a subset of key developmentally regulated genes in embryonic stem cells (ESC) (8, 9). The above scenario suggests that the aberrant de novo DNA methylation that so commonly affects cancer-related

genes could be a direct consequence of the underlying chromatin environment, driven by the presence of the polycomb-mediated mark H3K27me3 (5, 6, 10–13).

Here, we report the concurrent tumor-specific DNA methylation and silencing of a new set of neighboring genes, which are embedded in a stem cell-like bivalent chromatin structure. Our results, in concordance with recent reports (7), suggest a direct involvement of the two main cell-memory systems, PcG and Trx, in the aberrant setting of promoter DNA methylation. Moreover, we show that the presence of a bivalent chromatin landscape over large distances provides a chromatin-based explanation for the concurrent promoter DNA methylation of multiple neighboring genes (14, 15).

## Results

**Long Range Epigenetic Silencing Affects Multiple Genes in 5q35.2 and Is a Common Event in Colorectal Carcinogenesis.** To screen for recurrent DNA methylation changes in colorectal cancer, we used the Amplification of Unmethylated Alu (AUMA) method, which tracks unmethylated SmaI sites in Alu and other repetitive elements (16). Comparing normal and tumor tissue DNA with AUMA revealed that the Complexin 2 (CPLX2) gene CpG island is frequently hypermethylated (72%) in colon cancer cells [supporting information (SI) Fig. S1].

To determine whether CPLX2 hypermethylation was embedded in a larger region of epigenetic silencing, as previously described for the chromosomal region 2q14.2 (14), we performed an extended DNA methylation analysis of neighboring CpG islands in a region of 1.25 Mb in a subset of 5 normal-tumor pairs and seven colon cancer cell lines (Fig. 1A and Fig. S2). CPLX2 is located at 5q35.2, and the closest upstream CpG island, CpG72, associated with the HRH2 gene was found unmethylated in normal tissue but hypermethylated in all tumors and cell lines. Similarly, CpG145, and CpG79, associated with the DRD1 and SNCB gene promoter regions, respectively, were unmethylated in normal tissues and methylated in one or more tumors and most cell lines. Other neighboring CpG islands including CpG91, CpG114 (GPRIN1), CpG125 (AK124.837), and CpG78 (BC101340) were heavily methylated in most tumors and cell lines, although a moderate level of methylation was also observed in normal tissues. Finally, CpG islands associated with the genes SFXN1, THOC3, CLTB, RNF44 and TSPAN17 were maintained in a completely demethylated state in both normal and tumor cells (Fig. 1A).

The methylation status of the most recurrently methylated genes

Author contributions: J.R., C.G.F., M.G., and M.A.P. designed research; J.R., M.M., and L.V. performed research; J.R., M.G., and M.A.P. analyzed data; and J.R., M.G., and M.A.P. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

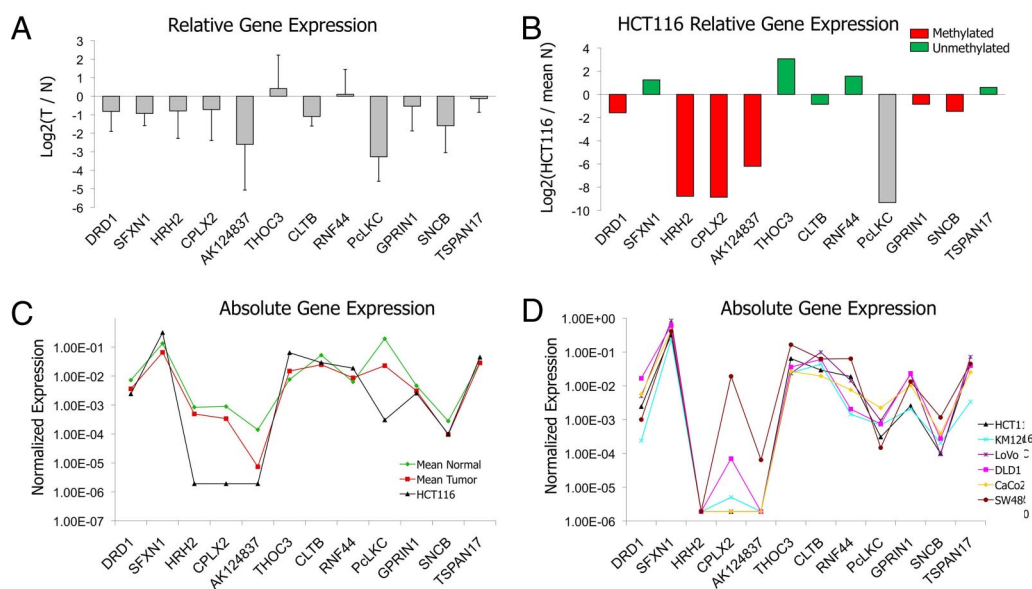
<sup>1</sup>To whom correspondence may be addressed. E-mail: markg@fhrc.org.

<sup>2</sup>To whom correspondence may be addressed at: Institut de Medicina Predictiva i Personalitzada del Càncer (IMPPC), Ctra Can Ruti, Camí de les Escoles, Badalona, 08916 Barcelona, Spain. E-mail: map@imppc.org.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0810133105/DCSupplemental](http://www.pnas.org/cgi/content/full/0810133105/DCSupplemental).

© 2008 by The National Academy of Sciences of the USA





**Fig. 2.** Gene expression and promoter DNA methylation status of 12 genes across 5q35.2 in normal-tumor pairs and colon cancer cell lines. (A) Mean expression values ( $\log_2$  of the tumor/normal ratio) of a series of 16 normal-tumor pairs. Error bars indicate standard deviation. Most genes were down-regulated in tumor samples. Individual data are illustrated in Fig. S3. (B) Relative expression of the 12 genes in HCT116 cells respectively, compared with the mean of 16 normal colonic mucosa samples. Bar color represents the methylation status of the respective promoter in HCT116 cells, except for PclKC, that does not have a promoter CpG island. (C) Mean absolute expression values across the 12 genes analyzed for the 16 normal tissues (green), the 16 tumor tissues (red), and the HCT116 cell line (black). (D) Absolute gene expression values in six colon cancer cell lines.

suggesting that drug treatment might also affect locus structure and organization.

A global enrichment for the AcH3K9 mark after 5-AzaC/TSA treatment was observed at all promoters tested, but particularly those of the DNA methylated genes. Increase of the AcH3K9 mark after 5-AzaC treatment was seen only over the genes HRH2, CPLX2, and SNCB (Fig. 5) exhibiting DNA methylated promoters in cancer cells. Unexpectedly, the AcH4K16 modification was slightly depleted after treatment with 5-AzaC alone at virtually all of the regions analyzed. In contrast, 5-AzaC/TSA treatment had the opposite effect, enriching for this modification over the down-regulated genes DRD1, HRH2, CPLX2, AK124.837, CLTB, PclKC, and SNCB, and also over the active TSPAN17 gene (Fig. 5). The highly expressed genes SFXN1, THOC3, and RNF44 showed very low or no enrichment of AcH4K16. Consistent with transcriptional down-regulation, highly expressed genes SFXN1 and RNF44 had their H3K4me3 levels reduced after the treatments, while a moderate depletion for this mark was seen only after 5-azaC/TSA treatment at the THOC3 and TSPAN17 promoters. On the other hand, silenced genes regardless of their DNA methylation status were enriched to a different degree for H3K4me3, and this was the only active mark that displayed a clear enrichment over the promoters of silenced genes after both 5-AzaC and 5-AzaC/TSA treatments.

Finally, whereas H3K27me3 levels were moderately enriched by 5-AzaC treatment, cotreatment with TSA clearly induced a global gain of this histone modification across the entire region (Fig. 5). Enrichment for this mark affected all genes, even though it was more clearly seen at the promoters of active genes that had been described previously to have very low or undetectable levels of this mark (Fig. 3).

**Components of Polycomb Repressor Complexes Are Found at the Inactive Genes.** The specific presence of H3K27me3 indicates that components of the polycomb group of proteins may be mediating, at least in part, the silencing that affects most of the genes across this region. To test this, we surveyed for the presence of the protein that catalyzes the deposition of the H3K27me3 mark, Enhancer of Zeste Homolog 2 (EZH2), identified in the context of different polycomb repressor complexes (PRC), and the protein BMI1, member of the polycomb repressor complex 1 (PRC1) that mediates the recognition of the H3K27 methylation mark (23). We show that both EZH2 and BMI1 are enriched over the promoters of the silenced genes where we had previously detected high levels of the

H3K27me3 mark (Fig. S7). The highest levels of EZH2 and BMI1 were detected over the promoters of the HRH2, CPLX2 and AK124.837 genes, whereas moderate levels were detected over the SNCB and PclKC genes. The DRD1 gene was an exception, displaying low levels of EZH2 and BMI1 despite having a bivalent promoter. DRD1 is the only gene in the region that is moderately expressed (Fig. 2) even though it displays dense promoter DNA methylation in the HCT116 cell line (Fig. 1 and Fig. S2). Additionally, specific enrichment of AcH3K16 after drug-induced reactivation over the promoter of silenced genes predicted the presence of the HDAC SIRT1 (24). Consistent with the results obtained for EZH2 and BMI1, SIRT1 was seen enriched over the promoter regions of the same genes (Fig. S7).

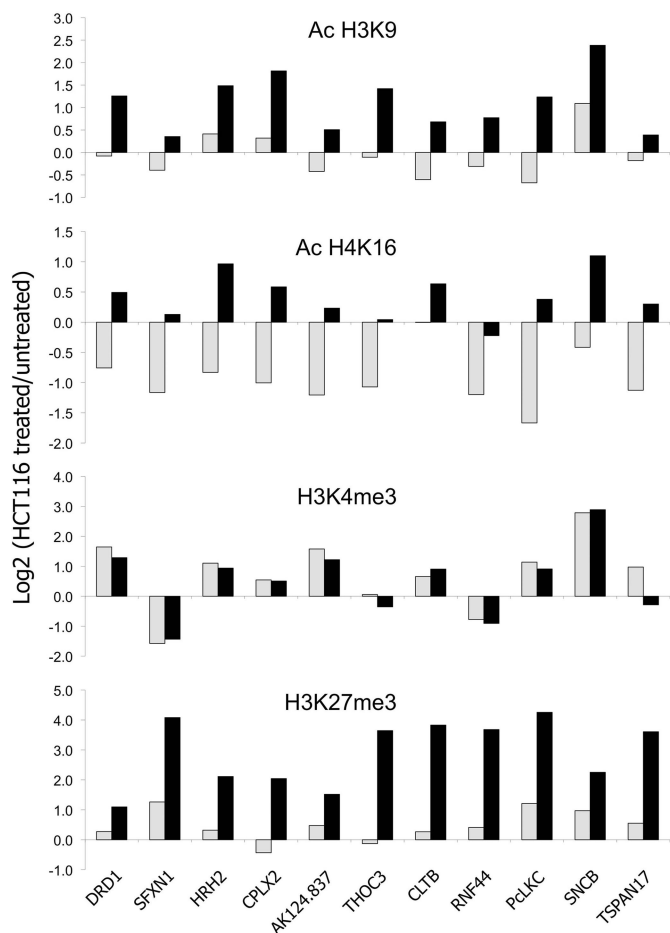
Both 5-AzaC and 5-AzaC/TSA treatments resulted in an increase in the EZH2 protein over the promoter regions of most genes (Fig. S7), in concordance with the enrichment of the H3K27me3 mark seen across 5q35.2. Opposite to EZH2, BMI1 was consistently depleted in all genes after drug treatment, particularly in cells treated with 5-azaC/TSA. Finally, whereas 5-azaC treatment alone increased the levels of SIRT1 at some promoters (HRH2, CPLX2, AK124.837 and CLTB), 5-AzaC/TSA treatment globally reduced the levels of this HDAC at all promoters tested, consistent with an increase of the AcH4K16 mark over the promoters of down-regulated genes after 5AzaC/TSA treatments (Fig. 5).

## Discussion

In the present work we provide another example of long range epigenetic silencing (LRES) affecting genes in the genomic region 5q35.2 in colorectal cancer, characterized by concurrent DNA hypermethylation together with global down-regulation of most of the genes in the region, none of which had been previously found to be hypermethylated in cancer. A similar phenomenon has been identified in chromosomal regions 2q14.2 (14) and 3p22.2 (15) in colorectal cancer. Confirming previous studies (14, 22), DNA methylation appears to play a dominant role in the silencing mechanism, as these genes cannot be reactivated by using TSA, unless first demethylated. Moreover, the inhibitory effects of DNA methylation extend over genes without a CpG island, such as PclKC, which can be reactivated after 5-azaC treatment, as we have previously reported for other genes in the region 2q14.2 (14).

In combination, the transcriptional, chromatin and CTCF analyses across this 1.2 Mb region provide evidence for the presence of isolated expression domains characterized by specific patterns





**Fig. 5.** Chromatin remodeling across 5q35.2 after drug treatment in the HCT116 cell line. The active histone modifications AcH3K9, AcH4K16, H3K4me3, and the inactive histone modification H3K27me3 in the drug-treated cells are shown as relative to the levels of the untreated cell line. Light gray bars correspond to the 5-AzaC treatment, and black bars correspond to the 5-AzaC/TSA treatment.

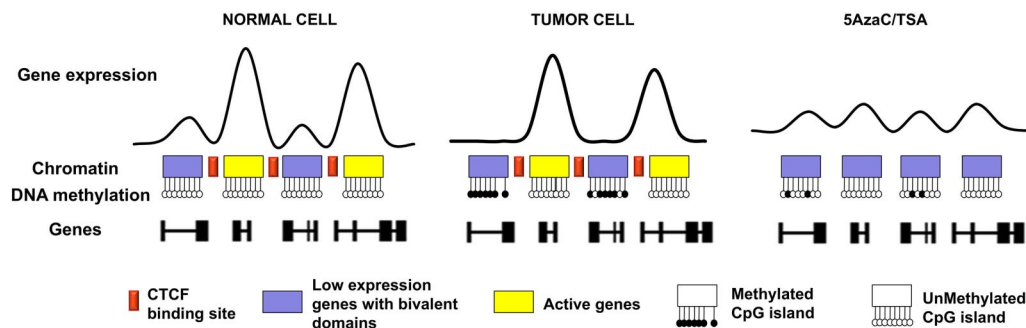
reported that bivalent chromatin is not limited to ESC: progenitor and terminal neurons also exhibit bivalent domains in specific polycomb targets (12, 13). Interestingly, when the bivalent state is resolved during cell differentiation, the active mark H3K4me2 is lost and de novo DNA methylation locks genes in a silent state (12). Here, we observe that DNA methylation is compatible with the retention of the bivalent chromatin state in cancer cells (Fig. 3),

suggesting that DNA hypermethylation might override the active chromatin mark H3K4me3, becoming a primary silencing mechanism in cancer cells.

Genome-wide analysis has demonstrated specific functional and structural features, including clustering in chromosomes, of the genes methylated in cancer, leading to the postulate that EZH2 may mark these genes as preferred substrates for targeted methylation (5, 6, 28). Our findings demonstrate that the concurrent DNA methylation of neighboring genes in the 5q35.2 region is associated with the presence of bivalent promoters in these genes, supporting the hypothesis that bivalent domains may favor concurrent DNA methylation and long range epigenetic silencing of clustered genes.

Drug-induced reexpression of DNA methylated genes in 5q35.2 is accompanied by an increase in the H3K9 and H4K16 acetylation levels, together with an increase in both H3K4 and H3K27 trimethylation levels (Fig. 5). The enrichment of the inactive mark H3K27me3 over the promoters of drug-reactivated genes has been also recently noted in double-knockout HCT116 cells lacking DNA methylation (7). Although we cannot exclude a passive accumulation of these two marks because of reduced cell division induced by drugs, it can be also suggested that cellular memory systems prevent these bivalent domains from being resolved into a more active configuration during drug-induced gene reactivation by balancing the levels of active H3K4me3 and inactive H3K27me3 marks. Moreover, we provide evidence that expression domains are disrupted upon 5-azaC and 5-azaC/TSA treatments, as denoted by the loss of CTCF insulator binding (Fig. 4) and the spreading of the H3K27me3 mark over the promoter regions of active genes (Fig. 5), which also results in their transcriptional down-regulation. This occurs in the presence of an increase in the levels of the AcH3K9 mark at all promoters after drug treatment. In this context, we propose a simplified model (Fig. 6) in which domain organization is relaxed upon drug treatment and profiles of interspersed active and silent chromatin are smoothed. The relaxation of the domain boundaries does not necessarily imply changes in gene activity but would facilitate the remodeling of chromatin by altering the accessibility to modulating factors. In this new context, the reduced expression of active genes despite the rise of active marks across the promoters in 5q35.2 would be consistent with a hypothetical dominant role of H3K27me3 mark over the rest of histone modifications. On the other hand, the retention of the bivalent domains over the reexpressed genes might prevent them from being highly transcribed after drug treatment, in concordance with the low transcription rates of these types of genes in ESC (8, 13). The presence of the PRC in the promoter regions of silenced genes along 5q35.2 is consistent with the presence of trimethylated lysine 27 over the same genes and the observation that genes frequently hypermethylated in tumors tend to be occupied by proteins of the PRC in human ESC (11). Upon AZA/TSA treatment, EZH2 levels are increased, paralleling H3K27me3 changes and allowing the

**Fig. 6.** Epigenetic model for the transcriptional down-regulation and drug induced epigenetic remodeling of genes across 5q35.2. In a normal cell setting, the promoters of all genes are unmethylated (white dots) and the presence of CTCF (red box) maintains the different chromatin and expression domains isolated. During carcinogenesis, genes with bivalent domains and kept at low transcription rates become DNA-methylated (black dots), stably silenced and insensitive



to differentiation or anti-proliferation cues (10, 11). This likely occurs through the recruitment of DNMT activities mediated by EZH2 (6), while active genes remain DNA methylation-free and can become up-regulated. Drug reactivation of the silenced genes not only fails to erase bivalent domains, but also induces the scattering of the silencing mark (H3K27me3) through the loss of CTCF-dependent insulation. Reactivation of silenced genes is further accompanied by an increase in the levels of active histone modifications and the partial loss of the proteins that mediate transcriptional silence.

retention of bivalent domains in promoters originally methylated. On the other hand, partial depletion of BMI1 is consistent with a moderate decrease of the polycomb complex and transcriptional reactivation of these genes. Additionally, the partial reduction of SIRT1 might explain the increase in the levels of Ach4K16 over the reexpressed genes, favoring the formation of a more transcription-competent environment over the genes with bivalent domains. These results are of particular relevance, because SIRT1, which participates in the maintenance of a stem cell state (29) in the context of the polycomb repressor complex 4 (PRC4) (30), is over-expressed in colon cancers, and its inhibition can lead to gene reactivation in the absence of DNA demethylation (24).

Different studies have identified stem cell-like signatures in cancer cells and have suggested that these molecular profiles contribute to defining the biological properties of tumor cells (7, 31–33). It has been suggested that tumor-specific de novo DNA methylation could be the result of an epigenetic program that functions in embryonic stages (5). Here, we have focused on a defined genomic region, where a new set of genes with cancer-specific DNA methylation cluster. Our findings emphasize the importance of studying the relationship between DNA methylation and chromatin signatures in the context of chromatin domains rather than isolated genes, as we provide evidence that the tran-

scriptional output for a given gene after drug treatment depends not only on its basal chromatin signatures, but also on the chromatin signatures of the neighboring genes.

Our results strengthen the concept that the cancer cell phenotype is sustained by cellular memory mechanisms directly involved in the maintenance of a bivalent chromatin state. The dominant nature of these bivalent epigenetic signatures has critical implications for understanding how cancer cells escape anti-proliferative and cell differentiation cues. The mechanisms regulating these epigenetic landscapes are likely to play a pivotal role in early stages of malignancy and endure the genotypic and phenotypic changes underlying tumor progression. Moreover, our results shed light on the complex effects of epigenetic therapies and may have important implications in the design and management of these treatments.

## Materials and Methods

Full details are in the SI Materials and Methods.

**ACKNOWLEDGMENTS.** We thank Dirk Schübeler for comments on the manuscript and helpful suggestions and Harvey Evans for help in revising the manuscript. J.R. was supported by a Generalitat de Catalunya fellowship and a Boehringer Ingelheim travel grant. This work was supported by Spanish Ministry of Science and Innovation Grant SAF2006/351, Consolider-Ingenio Grant 2010 CSD2006/49, and National Institutes of Health Grants R01 HL65440 and R37 DK44746.

1. Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell* 128:683–692.
2. Esteller M (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 8:286–298.
3. Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* 128:707–719.
4. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74:481–514.
5. Schlesinger Y, et al. (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat Genet* 39:232–236.
6. Vire E, et al. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439:871–874.
7. McGarvey KM, et al. (2008) Defining a chromatin pattern that characterizes DNA-hypermethylated genes in colon cancer cells. *Cancer Res* 68:5753–5759.
8. Bernstein BE, et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125:315–326.
9. Azuara V, et al. (2006) Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8:532–538.
10. Ohm JE, et al. (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat Genet* 39:237–242.
11. Widschwendter M, et al. (2007) Epigenetic stem cell signature in cancer. *Nat Genet* 39:157–158.
12. Mohn F, et al. (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 30:755–766.
13. Meissner A, et al. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766–770.
14. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark S (2006) Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nat Genet* 38:540–549.
15. Hitchins MP, et al. (2007) Epigenetic inactivation of a cluster of genes flanking MLH1 in microsatellite-unstable colorectal cancer. *Cancer Res* 67:9107–9116.
16. Rodriguez J, et al. (2008) Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells. *Nucleic Acids Res* 36:770–784.
17. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130:77–88.
18. Delaval K, Feil R (2004) Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 14:188–195.
19. Bao L, Zhou M, Cui Y (2008) CTCFBSDB: A CTCF-binding site database for characterization of vertebrate genomic insulators. *Nucleic Acids Res* 36:D83–87.
20. Barski A, et al. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837.
21. Kim TH, et al. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128:1231–1245.
22. Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat Genet* 21:103–107.
23. Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* 38:413–443.
24. Pruitt K, et al. (2006) Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation. *PLoS Genet* 2:e40.
25. Mikkelsen TS, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448:553–560.
26. Pan G, et al. (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 1:299–312.
27. Zhao XD, et al. (2007) Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* 1:286–298.
28. Keshet I, et al. (2006) Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat Genet* 38:149–153.
29. Vaquero A, Sternglanz R, Reinberg D (2007) NAD<sup>+</sup>-dependent deacetylation of H4 lysine 16 by class III HDACs. *Oncogene* 26:5505–5520.
30. Kuzmichev A, et al. (2005) Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. *Proc Natl Acad Sci USA* 102:1859–1864.
31. Clarke MF, Fuller M (2006) Stem cells and cancer: Two faces of eve. *Cell* 124:1111–1115.
32. Ben-Porath I, et al. (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* 40:499–507.
33. Spivakov M, Fisher AG (2007) Epigenetic signatures of stem-cell identity. *Nat Rev Genet* 8:263–271.