Linking DNA Methyltransferases to Epigenetic Marks and Nucleosome Structure Genome-wide in Human Tumor Cells

Citation: Jin, Bilian, Jason Ernst, Rochelle L. Tiedemann, Hongyan Xu, Suhas Sureshchandra, Manolis Kellis, Stephen Dalton, Chen Liu, Jeong-Hyeon Choi, and Keith D. Robertson. "Linking DNA Methyltransferases to Epigenetic Marks and Nucleosome Structure Genome-Wide in Human Tumor Cells." Cell Reports 2, no. 5 (November 2012): 1411–1424.

As Published: http://dx.doi.org/10.1016/j.celrep.2012.10.017

Publisher: Elsevier

Persistent URL: http://hdl.handle.net/1721.1/87008

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Cell Reports
Resource

Linking DNA Methyltransferases to Epigenetic Marks and Nucleosome Structure Genome-wide in Human Tumor Cells

Bilian Jin,[1,7] Jason Ernst,[3,4] Rochelle L. Tiedemann,[1] Hongyan Xu,[1,2] Suhas Sureshchandra,[1,2] Manolis Kellis,[3,4] Stephen Dalton,[5] Chen Liu,[6] Jeong-Hyeon Choi,[1,2,*] and Keith D. Robertson[1,*]
[1]Cancer Research Center
[2]Department of Biostatistics
Georgia Health Sciences University, Augusta, GA 30912, USA
[3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[4]MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA 02139, USA
[5]Paul D. Coverdell Center for Biomedical and Health Sciences, University of Georgia, Athens, GA 30602, USA
[6]Department of Pathology, Immunology and Laboratory Medicine, University of Florida, 1600 S.W. Archer Road, Gainesville, FL 32610, USA
[7]Present address: Institute of Cancer Stem Cells, Dalian Medical University Cancer Center, Dalian, Liaoning 116044, China
*Correspondence: jechoi@georgiahealth.edu (J.-H.C.), krobertson@georgiahealth.edu (K.D.R.)
http://dx.doi.org/10.1016/j.celrep.2012.10.017

SUMMARY

DNA methylation, mediated by the combined action of three DNA methyltransferases (DNMT1, DNMT3A, and DNMT3B), is essential for mammalian development and is a major contributor to cellular transformation. To elucidate how DNA methylation is targeted, we mapped the genome-wide localization of all DNMTs and methylation, and examined the relationships among these markers, histone modifications, and nucleosome structure in a pluripotent human tumor cell line in its undifferentiated and differentiated states. Our findings reveal a strong link between DNMTs and transcribed loci, and that DNA methylation is not a simple sum of DNMT localization patterns. A comparison of the epigenomes of normal and cancerous stem cells, and pluripotent and differentiated states shows that the presence of at least two DNMTs is strongly associated with loci targeted for DNA hypermethylation. Taken together, these results shed important light on the determinants of DNA methylation and how it may become disrupted in cancer cells.

INTRODUCTION

Methylation at the 5-position of cytosine within the CpG dinucleotide is an epigenetic modification of DNA that is linked to transcriptional repression when it is present within promoter regions, or to transcriptional activity when it is present within gene bodies (Jones, 1999; Maunakea et al., 2010). Global genomic DNA methylation patterns are established and maintained by the combined action of three enzymatically active DNA methyltransferases: DNMT1, DNMT3A, and DNMT3B. DNMT1 is traditionally referred to as the maintenance enzyme because it copies methylation after replication, whereas DNMT3A and DNMT3B are de novo enzymes that establish new patterns of methylation during differentiation (Baylin and Jones, 2011; Goll and Bestor, 2005). Although it has been extensively studied, much remains unknown regarding the division of labor among DNMT family members in methylating the genome.

Although there is evidence for redundancy among the DNMTs, there is also mounting evidence for functional specificity. During development, for example, genetic inactivation of *Dnmt3a* results in mice that develop to term and are grossly normal at birth but die by 4 weeks of age. In contrast, *Dnmt3b* null mice fail to develop normally past embryonic day 9.5 (E9.5) and no viable offspring are produced (Okano et al., 1999). In disease, genetic reduction of Dnmt1 levels reduced polyp formation in the Apc^{Min} murine colorectal cancer model (Laird et al., 1995), overexpression of Dnmt3b in the same system resulted in enhanced size and number of tumors (Linhart et al., 2007), and conditional knockout of *Dnmt3a* in a mutant K-ras lung cancer model promoted tumor growth and progression (Gao et al., 2011). In humans, mutations in *DNMT3B* are responsible for the majority of Immunodeficiency, Centromere Instability, and Facial Anomalies (ICF) syndrome cases (Ehrlich et al., 2006). Collectively, these findings emphasize that DNMT family members possess distinct nonoverlapping functions; however, the regions of the genome that are bound and regulated by each DNMT, and the mechanisms by which they are directed to these sites, remain largely unknown.

One process that plays a role in regulating methylation patterns is the extensive array of posttranslational modifications of amino acid side chains on the core histones H2A, H2B, H3, and H4, which collectively form the nucleosome core particle and package ~150 bp of DNA. Histone modifications, like DNA methylation, play key roles in transcriptional regulation, genome stability, and embryonic development (Berger, 2007). Histone acetylation is generally associated with transcriptional activation, whereas the relationship between histone lysine methylation and transcriptional activity is dependent on the position and state (mono-, di-, or trimethylation [or me1, me2, and me3, respectively]) of the modification (Barski et al., 2007; Wang

et al., 2008). H3K27me3 and H2AK119 monoubiquitination (H2AK119Ub), transcriptionally repressive marks mediated by the PRC2 and PRC1 Polycomb complexes, respectively, have been positively associated with DNA methylation (Kallin et al., 2009; Viré et al., 2006). Loci regulated by Polycomb in normal cells display an increased propensity to acquire aberrant DNA hypermethylation in cancer (Jin et al., 2009; Ohm et al., 2007; Schlesinger et al., 2007). The underlying positioning of nucleosomes themselves also influences DNA methylation (Chodavarapu et al., 2010). Taken together, these results strongly suggest that histone modifications and nucleosomal structure play key roles in determining DNA methylation patterns; however, exactly how these chromatin determinants differentially affect each DNMT remains to be elucidated.

In this study we sought to gain a more comprehensive understanding of the unique, cooperative, and overlapping targets of each enzymatically active DNMT, and to define the relationships among these patterns and DNA methylation histone modifications, and nucleosomal structure in human cells. Using pluripotent human embryonic carcinoma (hEC) cells in both undifferentiated (UD) and differentiated (DF) states as a model system, we mapped the genome-wide localization of key histone modifications positively and negatively associated with DNA methylation using chromatin immunoprecipitation sequencing (ChIP-seq). We established global DNA methylation patterns using methyl CpG-binding domain protein-based enrichment coupled with next-generation sequencing (NGS; MBD-seq), and defined DNMT1, DNMT3A, and DNMT3B localization using ChIP-seq. Finally, we employed micrococcal nuclease digestion-based whole-genome sequencing (MNase-seq) to establish nucleosome positioning. We then used the global patterns of epigenetic marks and transcription to examine features of regions bound by each DNMT before and after differentiation. This work establishes direct connections between the DNMTs and other facets of the epigenome, and provides important information about how DNA methylation is targeted.

RESULTS

Characterization of Global Histone Modification Patterns in hEC Cells

To characterize the EC epigenome and how differentiation alters epigenetic marks, we treated NCCIT hEC cells with 10 μM all-*trans* retinoic acid (RA) or left them untreated, and then prepared formaldehyde-fixed chromatin. NCCIT cells are derived from a nonseminomatous germ cell tumor and possess features of both seminomas and embryonal carcinomas. They resemble embryonic stem (ES) cells in a number of ways, but are adapted to grow as tumors. Upon treatment with RA, NCCIT cells differentiate into derivatives of all three embryonic germ layers and extraembryonic lineages (Sperger et al., 2003; Figure S1A). We found that after RA treatment for 21 days, expression of DNMT3B decreased, whereas DNMT1 and DNMT3A levels remained relatively constant (Figures S1B–S1C). Global expression profiling revealed a high degree of correlation between NCCIT and WA09 (H9) human ES (hES) cells (Figure S1D), and ontology analysis showed marked overrepresentation of terms related to embryogenesis and development for RA-DF NCCIT cells (Figure S1E).

Based on these data, we focused subsequent RA treatments on day 7 as a balance between inducing the differentiation program while retaining sufficient levels of DNMT3B to perform subsequent ChIP-seq experiments. Antibodies against the activating mark H3K4me3; the repressive marks H3K9me3, H3K27me3, and H2AK119Ub; and the gene-body-associated mark H3K36me3 were used for ChIP-seq (summarized in Table S1). The ChIP-seq data were annotated and the association of each histone modification with intragenic features was examined (summarized in Figure 1A). H3K4me3 was most enriched in promoter regions, and H3K27me3 and H2AK119Ub were also high at the transcription start site (TSS). H3K36me3, in contrast, was most highly enriched in gene bodies, and gene 3′-ends were mildly enriched for several histone marks (Figure 1A). Given the importance of CpG islands (CGIs) in the biology of DNA methylation we also stratified our ChIP-seq data to show marks associated with CGIs in different contexts. In general, CGIs associated with genes were marked by H3K4me3 and nearly devoid of H3K9me3 (Figure 1B). Histone modifications within a portion of the *HOXA* locus are shown in Figure 1C to illustrate a representative bivalent gene (H3K4me3, H3K27me3, and H2AK119Ub-positive) transitioning to a monovalent (H3K4me3 only) state following RA-induced differentiation. This figure also highlights the enrichment of H3K36me3 in the gene body. Confirmation of the presence of select histone modifications mapped by ChIP-seq at loci relevant to development and RA-induced differentiation of NCCIT cells, using ChIP quantitative PCR (ChIP-qPCR), is shown in Figure S2.

We further investigated the localization of each histone mark by generating normalized tag density distribution plots across genes by type (protein coding versus noncoding; Figure 1D), stratified by promoter CpG density (Figure S3A), and across CGIs within different contexts (Figure S3B). H3K4me3 localized to a sharp double peak spanning the TSS with a trailing shoulder extending into the gene body (Figure 1D). H3K27me3 and H2AK119Ub profiles were similar in UD NCCIT cells. After differentiation, however, the H3K27me3 patterns changed markedly, and were almost a mirror image of their distribution in the UD state (Figure 1D). H3K36me3 was depleted over the TSS but was markedly and progressively enriched across gene bodies. We also examined the association of histone modifications with promoters of different CpG density. Stratifying genes in this way did not alter distribution profiles, but revealed that high CpG-density promoters (HCPs) and low CpG-density promoters (LCPs) were the most and least enriched, respectively, for most histone modifications (Figure S3A). CGIs were markedly enriched for H3K4me3, particularly those associated with promoters (Figure S3B). H3K27me3 and H2AK119Ub were similarly enriched at CGIs in the UD state; however, both marks were reduced after differentiation and H3K27me3 became depleted (Figure S3B), suggesting that many CGIs are bivalent in UD NCCIT cells but become monovalent after differentiation. Using recently published ChIP-seq data (Adamo et al., 2011), we compared histone marks in hES and NCCIT cells. There was marked overlap between genes bound by H3K4me3 and H3K27me3 in hES and UD NCCIT cells (Figure 1E); however, NCCIT cells had more H3K27me3-containing loci. This result also likely accounts for NCCIT cells possessing a greater

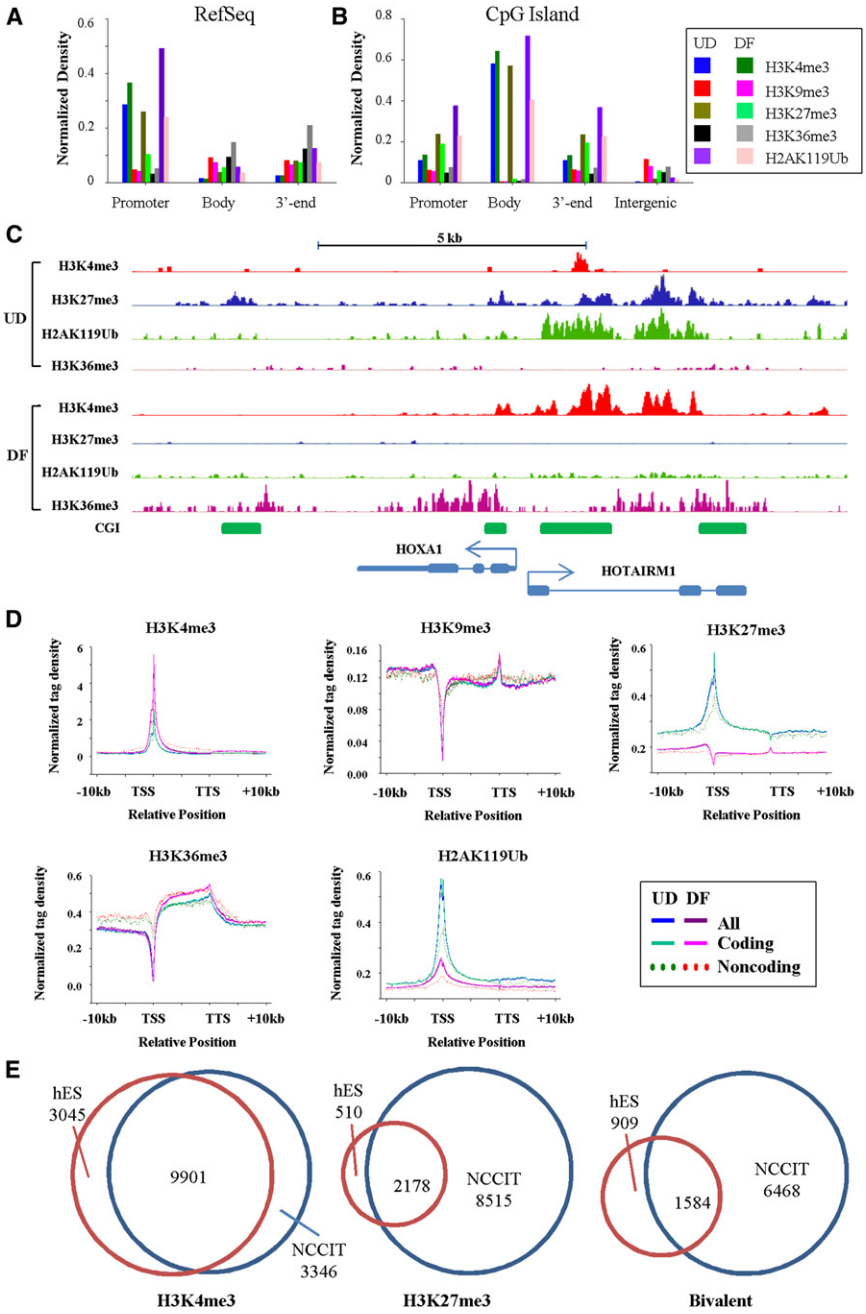

Figure 1. Global Analysis of Histone Modification Patterns in EC Cells

(A) Relationships between genomic features and ChIP-seq peaks for five histone modifications in UD and 7 day RA- DF NCCIT cells. The *y* axis represents the fraction (proportion) of the total length of the peaks in a genomic feature normalized to the total length of that feature present in the genome. Features were defined with hg19 RefSeq gene as promoter (TSS-1kb), gene body (TSS to TTS), and 3′-end (TTS+1kb).

(B) Enrichment of histone modifications in CGIs based on context as in (A).

(C) Representative ChIP-seq results at a locus that transitions from bivalent to monovalent during differentiation. Bent arrow: TSS (+1).

(D) Normalized tag distribution profiles of five histone modifications across intragenic regions before and after differentiation. Each gene body is normalized to 0%–100% from the TSS to the TTS, with 10 kb upstream (promoter) and 10 kb downstream of the gene body shown. Tag densities were normalized to total mapped read numbers in each sample. Different colors are used to represent UD or DF conditions, and genes are further stratified by protein coding potential.

(E) Venn diagrams summarizing the overlap between histone modifications and bivalent loci in UD NCCIT compared with hES cells.

See also Figures S1, S2, and S3, and Tables S1, S2, and S3.

number of bivalent loci than hESs. Taken together, these data reveal similarities between hES and hEC cells in terms of both expression and histone modification patterns, although notable differences also exist that may reflect their distinct cellular origins or state of transformation.

Mapping the Localization of DNMT1, DNMT3A, DNMT3B, and DNA Methylation Genome-wide in NCCIT Cells

To examine how each DNMT interfaces with DNA methylation and histone modifications, we used antibodies directed against DNMT1, DNMT3A, and DNMT3B (Figure S1C) in ChIP-seq experiments, and determined global DNA methylation patterns using methyl-CpG binding domain protein-based enrichment coupled with NGS (MBD-seq), essentially as described in Serre et al. (2010). Validation of the specificity of MBD-based enrichment for known methylated and unmethylated loci is shown in Figure S4A. DNMT ChIP-seq and MBD-seq data are summarized in Table S1. DNMT3A was least enriched at intragenic loci, whereas DNMT1 and particularly DNMT3B were present at higher levels in bodies and 3′-ends. After differentiation, DNMT3B levels generally declined in intragenic regions and CGIs (Figures 2A and 2B). Representative binding profiles are shown in Figure 2C. Because repetitive regions are established targets of DNA methylation we also stratified our localization data by repeat class, which showed that DNMT3B was enriched in SINEs, and DNMT3A was enriched in satellite repeats (Figure S4B). ChIP-qPCR and MBD pull-downs coupled to qPCR were used to confirm DNMT binding and DNA methylation respectively (Figure S2), showing the validity of our whole-genome-based analyses.

To further examine how each DNMT occupied intragenic regions, we created tag density plots across genes stratified by protein coding potential and CpG density, across CGIs, and

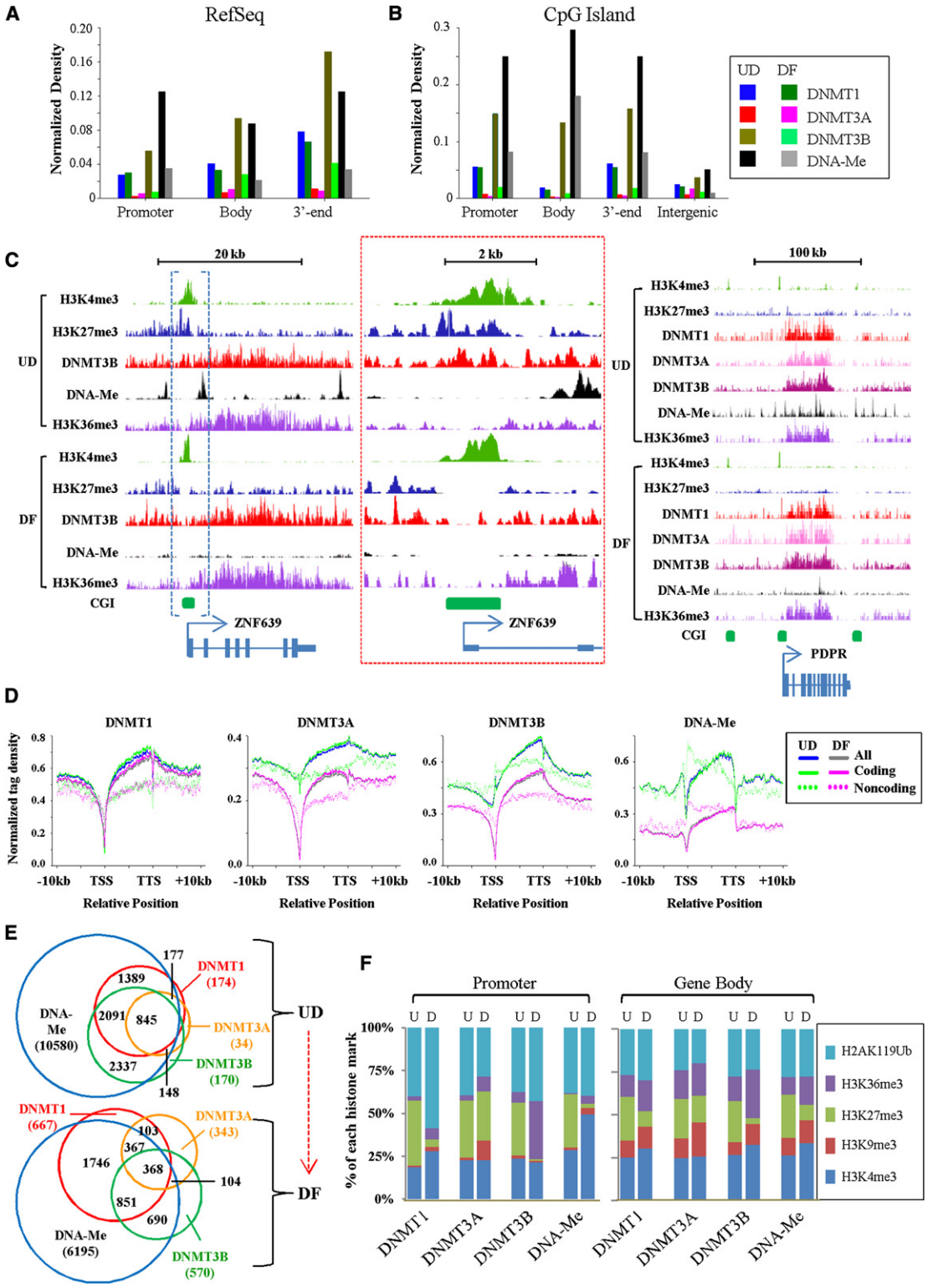

Figure 2. DNMT1, DNMT3A, DNMT3B, and DNA-Me Show Distinct and Overlapping Distributions across the NCCIT Genome

(A) Enrichment of each enzymatically active DNMT and DNA-Me within intragenic features of UD and RA- DF NCCIT cells as described in Figure 1.

(B) Enrichment of DNMTs and histone modifications within CGIs based on context.

(C) Representative ChIP-seq and MBD-seq results for the indicated DNMTs, DNA-Me, and histone modifications at two loci. In the left panel, the region enclosed by blue brackets is enlarged in the center (red boxed region), to highlight reduced DNMT3B binding at the TSS upon differentiation.

across exons. DNMT1 and DNMT3B were most enriched in gene bodies, and, like H3K36me3, increased from 5′ to 3′ (Figure 2D). DNMT1 was depleted at both the TSS and the transcription termination site (TTS). DNMT3A and DNMT3B were depleted at the TSS in DF cells, but were either similar to gene-body levels or slightly enriched at the TSS before differentiation. DNA methylation like DNMT1, was depleted at the TSS and TTS, but progressively elevated across the entire gene body (Figure 2D). The depletion of DNA methylation and DNMTs coincided precisely with the TSS-associated peak of H3K4me3 (Figure S5A); however, H3K4me3 trails into and overlaps with gene-body DNA methylation. Stratification by CpG density revealed that HCP genes generally had the highest levels of DNMT binding in the gene body and the largest depletion at the TSS. Interestingly, peaks of DNMT1, DNMT3B, and DNA methylation at the TSS distinguished the intermediate CpG-density promoter (ICP)/LCP from HCP genes (Figure S5A). DNMT1 and DNMT3A were depleted from CGIs; however, DNMT3B was enriched at CGIs before differentiation (Figure S5B). DNA methylation was also slightly elevated at CGIs overall, although promoter-associated CGIs were always the least methylated, and absolute levels of H3K4me3 were much greater than DNA methylation at CGIs (Figure S5B). This result suggests that DNMT3B is capable of binding and methylating a fraction of CGIs even though they contain H3K4me3. We plotted DNMT and DNA methylation patterns across exons as a function of coding potential and CpG density. With the exception of DNMT3A, all DNMTs and DNA methylation were at least mildly enriched in exons (Figure S5C). Interestingly, noncoding RNA genes were distinct in their profiles for many epigenetic marks (Figures 1D, 2D, and S5C), including DNMT localization. This finding could reflect the fact that noncoding transcripts generally are expressed at lower levels than coding transcripts (Birney et al., 2007), necessitating distinct epigenetic regulation. Taken together, our results show that DNMT1, DNMT3A, and DNMT3B have both unique and overlapping distributions across genomic features depending on cellular differentiation status.

Interrelationships among DNMTs, DNA Methylation and Histone Modifications

We further examined the overlap among the epigenetic marks. Venn diagrams focused on DNA methylation showed that although the majority of DNMT-bound regions were methylated, nearly half of the methylated loci did not appear to be stably occupied by a DNMT (Figure 2E). In addition, a small number of regions in UD cells were occupied by at least one DNMT, but were not positive for methylation, and this class of sequences increased >4-fold after differentiation (Figure 2E). Gene promoters and bodies were also separately compared, which revealed even more DNMT-bound unmethylated loci within gene bodies (Figure S5D). At promoters, DNMT1 overlapped the most with DNMT3B and methylation, and the least with DNMT3A. After differentiation, methylated promoters were most highly associated with DNMT1. Finally, we examined the co-occupancy of DNMT-bound loci with five different histone modifications. In UD NCCIT cells, promoter-associated DNMTs overlapped mostly with H3K27me3 and H2AK119Ub (Figure 2F). After differentiation, DNMT1 overlap with H2AK119Ub increased and DNMT3B overlap with H3K36me3 increased. In gene bodies, overlap of DNMTs with H3K9me3 and H3K36me3 increased after differentiation. Overlap between DNA methylation and H3K4me3 was due in large part to the TSS-associated H3K4me3 peak trailing into the body (as illustrated in Figure S5A). These results demonstrate that although most regions of the genome bound by DNMT are methylated, methylation is not a simple sum of DNMT binding patterns.

Differential Enrichment of DNMTs and Epigenetic Marks at Loci Targeted for DNA Hypermethylation—A Comparison of Two Systems

Given that both NCCIT EC and WA09 ES cells are pluripotent, but one is immortal (ES) and the other is transformed (EC), we were interested in comparing their patterns of DNA methylation to define chromatin-associated features of loci that sustain aberrant DNA methylation changes linked to the cancer phenotype. To accomplish this, we performed MBD-seq of WA09 hES cells to permit direct comparison of genome-wide methylation patterns using the same assay platform (Table S1). NCCIT and WA09 methylation patterns over the entire genome, at promoters, and in gene bodies were generally less well correlated with each other than gene expression patterns (Figure 3A, compare with Figure S1D). Tag density plots comparing hES and NCCIT intragenic DNA methylation reveal qualitatively similar patterns, although hES cells generally have higher levels of gene-body methylation and differ to some degree at the TSS in all CpG density classes (Figure 3B). Representative similarly and differentially methylated loci are shown in Figure 3C.

We hypothesized that loci differing in DNA methylation patterns between ES and EC represent the combined impact of adaptation to tumor growth and distinct cellular/developmental origin. Cell-culture-associated differences are unlikely to be a major contributor, because both lines have been extensively cultured. Consistent with the hypothesis that some methylation differences between ES and EC are the result of different developmental origins, an ontology analysis showed enrichment of processes related to meiosis as hypomethylated in EC relative to ES (Figure 3D). Because other methylation changes are likely linked to transformation, we examined the association of all differentially methylated loci with histone modifications, DNMT localization, and other characteristics of loci targeted for aberrant methylation, such as PcG regulation (Schlesinger et al., 2007). Furthermore, because many promoter CGIs are targeted

(D) Normalized distribution profiles of DNMTs and DNA-Me across intragenic regions in NCCIT cells before and after differentiation as described in Figure 1.
(E) Venn diagrams summarizing the overlap between genes marked by at least one DNMT and DNA-Me in UD and DF NCCIT cells. Further breakdown by promoter and body is shown in Figure S5D.
(F) Proportion of each histone mark overlapping with DNMT-bound or DNA methylated regions at promoters (left) or gene bodies (right). U, undifferentiated; D, RA-differentiated NCCIT.
See also Figures S1, S2, S4, and S5, and Tables S1, S2, and S3.

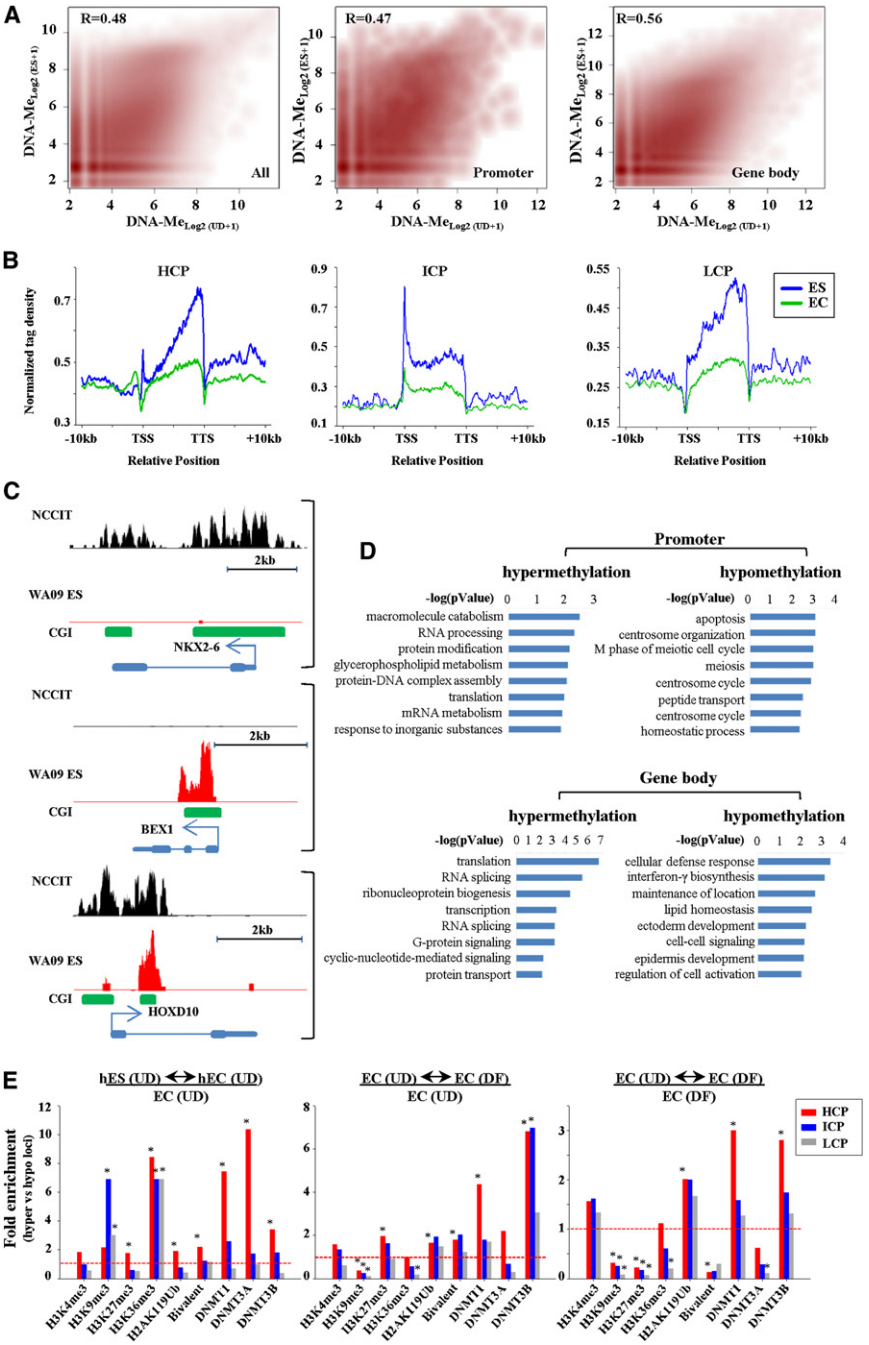

Figure 3. DNA-Me and DNMT Localization in Normal versus Malignant Stem Cells, and Features of Loci Targeted for Methylation

(A) MBD-seq data from pluripotent NCCIT and WA09 hES cells compared over the entire genome (left), promoters (middle), and gene bodies (right). Pearson correlation coefficients are shown.

(B) Normalized tag distribution profiles of DNA-Me across gene bodies in UD hES and NCCIT cells.

(C) Representative MBD-seq results in UD NCCIT (black) compared with WA09 hES cells (red) at loci that are hypomethylated (top, NKX2-6), hypermethylated (middle, BEX1), or similarly methylated (bottom, HOXD10) in WA09 relative to NCCIT.

(D) Biological processes enriched in differentially methylated genes in NCCIT (UD) compared with WA09 using DAVID.

(E) Left panel: Enrichment of DNMTs, histone modifications, and bivalent status at loci that are hypermethylated in EC (UD) relative to ES, compared with loci that are hypomethylated in EC relative to ES cells. Middle-right panels: Enrichment of DNMTs and histone marks at loci that become hypermethylated upon RA-induced differentiation of NCCIT cells, relative to loci that become hypomethylated upon differentiation ($\geq$4-fold change). Enrichments are relative to the histone modifications and DNMT binding patterns derived from NCCIT in the UD condition in the middle panel and the DF condition in the right panel. Significance was assessed using the chi-square test (*$p < 0.01$). Dashed line: 1X enrichment level.

for hypermethylation in adult cancers, we stratified all loci that become hypermethylated in UD NCCIT compared with hES by their promoter region CpG density. This analysis revealed a modest but significant enrichment of the PcG-mediated marks H3K27me3 and H2AK119Ub, both individually and coexistent (bivalent), and H3K36me3 at HCP genes that were hypermethylated in UD NCCIT compared with hES cells (Figure 3E, left graph). Others have shown that genes repressed by PcG in ES cells are 2–3 times more likely to be targeted for aberrant DNA hypermethylation in cancers than non-PcG-regulated loci (Schlesinger et al., 2007), consistent with our findings here. Genes with lower CpG density targeted for DNA hypermethylation showed marked enrichment for H3K9me3, consistent with a report that aberrantly hypermethylated loci in adult cancers display H3K9me3 enrichment in UD NTERA2 EC cells (Ohm et al., 2007). Interestingly, binding of all DNMTs, and particularly DNMT1 and DNMT3A, was highly enriched at HCP genes that were methylated in EC relative to ES (Figure 3E), suggesting that their presence primes loci for de novo methylation.

Although the ES-EC comparison is informative, it makes use of different cell lines, and mapping of DNMT binding in ES is not currently available. By using loci that become hypermethylated in DF compared with UD NCCIT cells, one can directly compare epigenetic marks involved in methylation targeting by comparing two states of the same cell line. To accomplish this, we correlated global DNA methylation patterns in UD and DF NCCIT to identify hyper- and hypomethylated loci (not shown). We then examined these genes for the presence of histone modifications and DNMTs that we mapped in both UD and DF states. As for the ES-UD comparison, loci that became hypermethylated

upon NCCIT differentiation showed modest but significant associations with both PcG-mediated marks and bivalent status at HCP genes (Figure 3E, middle graph). DNMT3B (at HCPs and ICPs) and, to a lesser extent, DNMT1 (at HCPs) were also highly enriched at loci in UD cells that sustained an increase in methylation upon RA treatment. In contrast to the ES-EC (UD) analysis, comparison of UD and DF NCCIT states revealed that hypermethylated loci were not enriched for DNMT3A and were significantly depleted for H3K9me3. The former difference may be attributed to the high expression of DNMT3B in EC cells, and the latter may be explained by H3K9me3 enrichment already being present on some genes in UD cells because they are transformed (Figure 3E, left panel; Ohm et al., 2007). Interestingly, if we examine chromatin modifications associated with hypermethylated loci once they become methylated (in the DF condition), many of the same marks remain significantly enriched (e.g., DNMT1 and DNMT3B) or depleted (e.g., H3K9me3); however, a significant depletion for both PcG regulated loci (defined by H3K27me3) and bivalency (Figure 3E, right graph) is now observed. This result suggests that some loci targeted for DNA hypermethylation undergo an epigenetic switch in repressive marking (H3K27me3 to DNA methylation) and/or change from a poised to a fully repressed state. Taken together, these data strongly suggest that DNMT binding does not necessarily equate to methylation of the bound region; however, the presence of at least two DNMTs (DNMT1 and DNMT3A or DNMT3B), and the presence/absence of other histone modifications (that vary to some extent by cell type) render loci more susceptible to, or target, transformation and/or differentiation-associated DNA hypermethylation events.

Relationships among DNA Methylation DNMT Occupancy, and Transcription

Because DNA methylation and each DNMT differentially associate with genes that undergo methylation changes based on their promoter CpG density (Figure 3), we first examined how CpG density relates to expression in NCCIT cells using our microarray data (Figure S1D). This analysis demonstrated a clear, direct correlation between expression and promoter CpG density, with the majority of highly expressed genes being in the high-CpG-density class (HCP), the majority of moderately expressed genes falling into the ICP class, and most lowly expressed genes grouping into the LCP class (Figure 4A). Given the associations between expression and promoter CpG density we observed, we stratified the genes into three combined classes (HCP-high expression, ICP-medium expression, and LCP-low expression) to examine the relationships among DNA methylation, DNMTs, and the three gene classes, as shown in Figures 4B–4E. These plots show that DNA methylation at the HCP-high class of genes exhibits a peak of enrichment followed by a region of depletion flanking the TSS (hill-valley), followed by a gradual increase across the gene body and another sharp reduction at the TTS (Figure 4B). For both UD and DF conditions, DNMT1 is markedly depleted at the TSS of HCP-high genes, elevated progressively across the gene body, followed by another depleted region at the TTS. Unlike DNA methylation DNMT1 decreases only gradually in the 3′ region (Figure 4C). DNMT3A shows the largest differentiation-specific differences among DNMTs, particularly at the TSS (Figure 4D). Finally, DNMT3B shows the greatest gene-body enrichment of all DNMTs, particularly in the HCP-high class (Figure 4E). A unique feature of DNMT3B in UD cells is a spike of enrichment just after the TSS in HCP-high genes. Tag density plots for these four marks stratified by expression level only are shown in Figure S6 for comparison. Genes were further stratified based on whether they were upregulated, downregulated, or displayed no differential expression after RA treatment, revealing that upregulated genes generally displayed lower levels of DNMTs and DNA methylation particularly in the UD state, than genes that became downregulated (Figure S6). Taken together, these data reveal that DNMT1 and DNMT3B are most tightly linked to gene bodies, and that their enrichment increases as a function of expression and CpG density, although the promoter region itself generally shows poor enrichment of these factors. These results also illustrate subtle but distinct localization patterns of each DNMT, particularly at the TSS and TTS, and that no single DNMT exactly parallels DNA methylation patterns.

Global Nucleosome Positioning in NCCIT Cells Reveals Links with DNMT Binding and DNA Methylation, and Distinct Differences from Normal Human Cells

Recent studies suggested a positive link between DNA methylation and nucleosome occupancy (Chodavarapu et al., 2010), but, to our knowledge, global nucleosome positioning in transformed human cells has not previously been reported. To examine these associations, we isolated mononucleosomes from UD and DF NCCIT cells by micrococcal nuclease (MNase) digestion. Size-fractionated, gel-purified DNA of ~150 bp was then subjected to massively parallel sequencing (MNase-seq; Figure S7A) as described previously for human T cells (Schones et al., 2008). We obtained 359,468,369 and 328,468,369 reads from UD and RA-DF NCCIT cells, respectively, resulting in 16–18X genome coverage by 147 bp nucleosome footprints (as calculated in Valouev et al., 2011), which is comparable in depth to a study using human T cells (Schones et al., 2008). The DANPOS algorithm was used to map nucleosomes, resulting in identification of 9,772,482 and 10,148,751 positioned nucleosomes in pluripotent and RA-DF NCCIT cells, respectively.

We then examined the relationship between nucleosome occupancy and gene features, promoter CpG density, and expression. To compare our data derived from a transformed cell line with nucleosome positions in normal human cells, we used a previously published CD4 T cell MNase-seq data set (Schones et al., 2008). Tag density plots across intragenic loci show that nucleosome occupancy is relatively constant over genes in NCCIT cells, being marked by a depleted region at the TSS in both the UD (Figure 5A) and DF states (Figure S7B). Nucleosome positions in CD4 T cells showed notable similarities and differences compared with NCCIT cells. For example, nucleosome density increased more across the gene body in CD4 T cells than across gene bodies in NCCIT cells (Figure 5A). Tag density plots centered on the TSS highlight the nucleosome-depleted region (NDR) present in both cell types. A TTS-centered plot emphasizes one of the most prominent differences: whereas CD4 T cells have another NDR at this location, NCCIT cells have a peak of nucleosome enrichment (Figure 5A). Stratifying genes

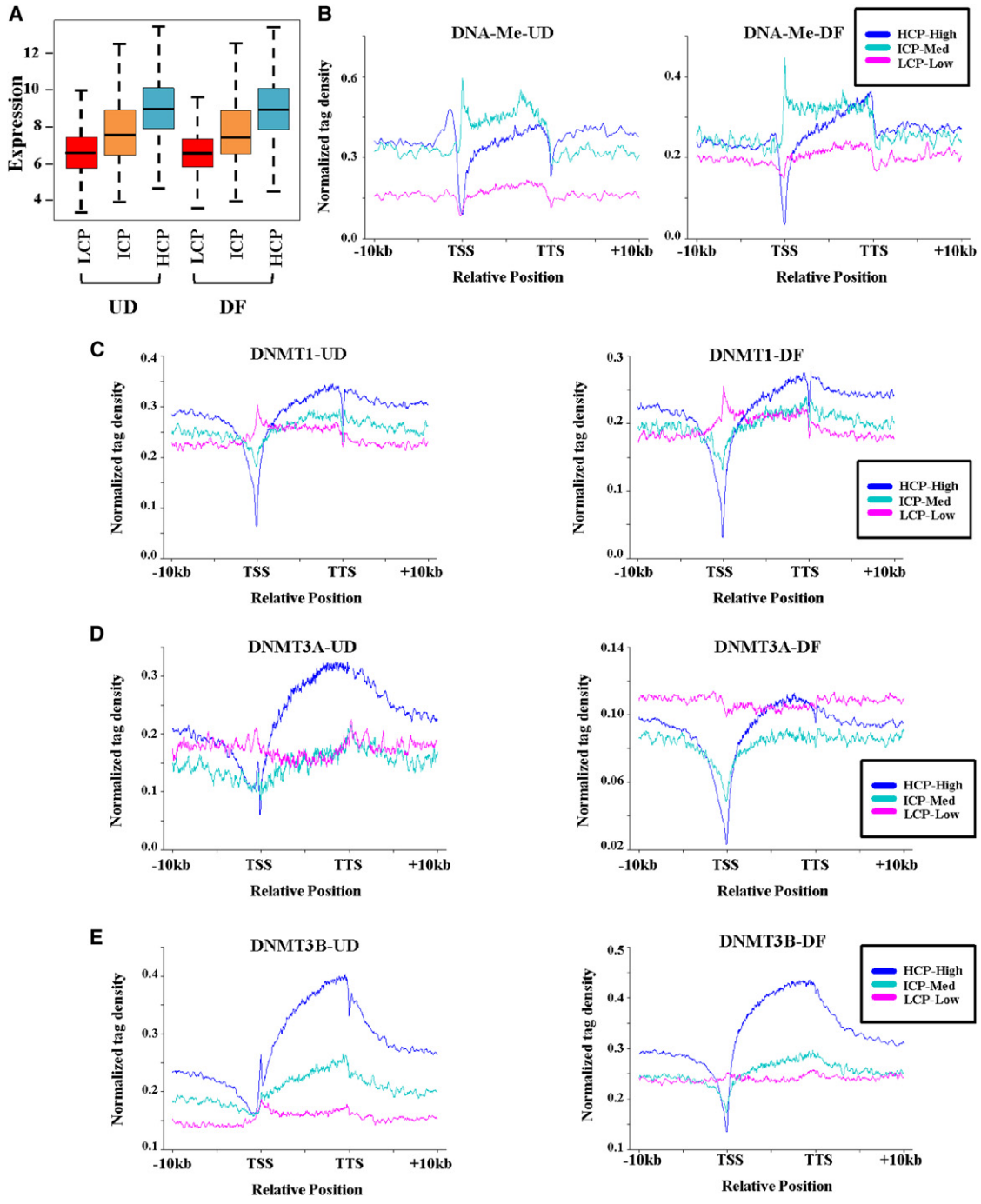

Figure 4. DNMTs Are Differentially Enriched in Gene Bodies as a Function of Expression and Promoter CpG Density

(A) Relationship between promoter CpG density and expression in NCCIT cells. The box plot shows the relationship between $\log_2$ transformed gene expression level (stratified as low expression < 7, medium expression 7–9, and high expression > 9) and promoter CpG densities (HCP, ICP, and LCP). The bottom and top of each box represent the lower and upper quartiles, respectively; the bar represents the median; and the whiskers represent the lowest and highest data within the 1.5 interquartile range.

(B–E) Tag density plots illustrating relationships between DNA-Me/DNMTs and expression stratified into HCP-high (blue), ICP-medium (green), and LCP-low (red) classes. Normalized distribution profiles of (B) DNA-Me, (C) DNMT1, (D) DNMT3A, and (E) DNMT3B in NCCIT cells before (left) and after (right) differentiation are shown.

See also Figure S6.

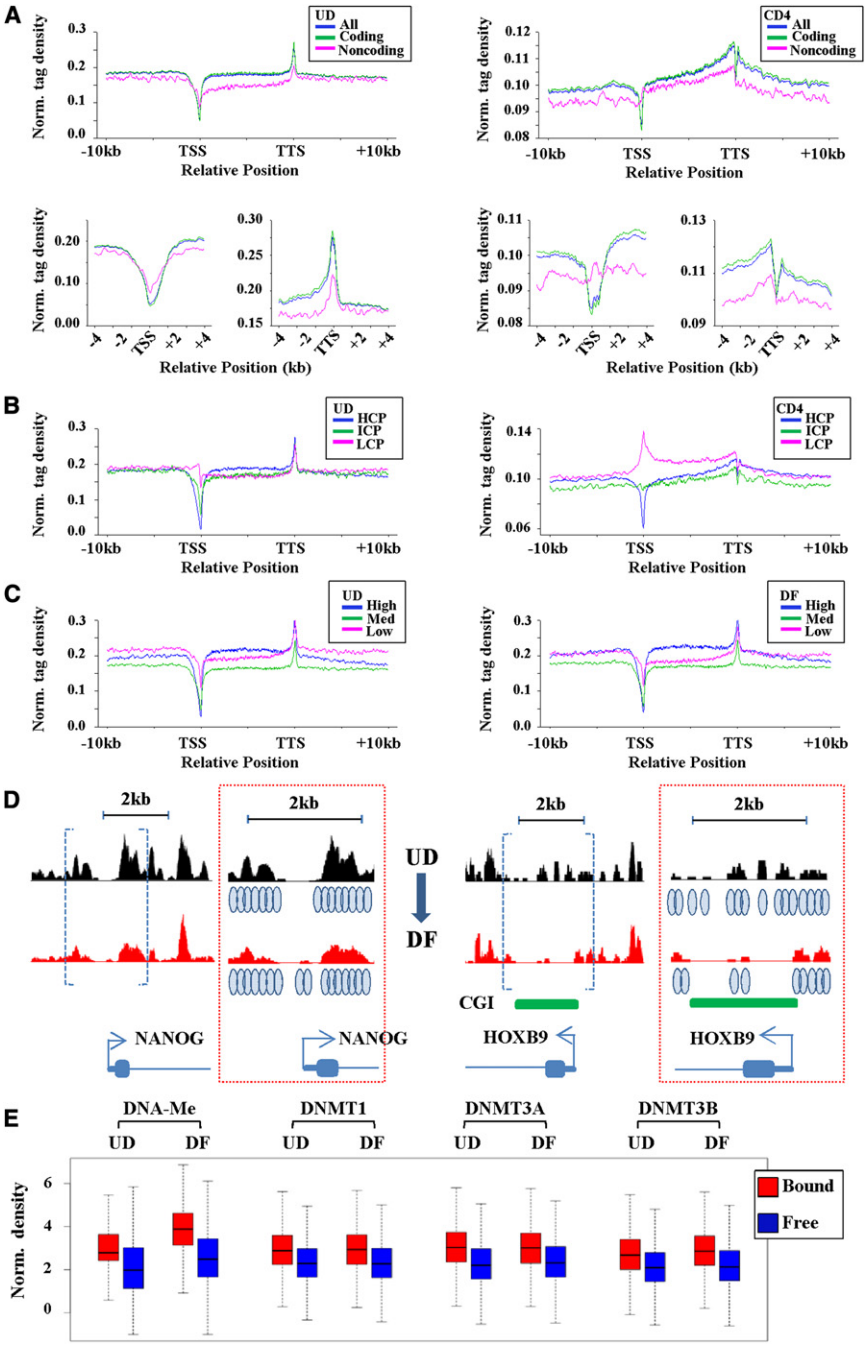

Figure 5. Relationships among Global Nucleosome Positioning, Genomic Features, Transcription, DNMTs, and DNA-Me

(A) Normalized distribution profiles for nucleosome binding across intragenic regions for all, protein coding, and nonprotein coding genes in pluripotent NCCIT cells. MNase-seq data from human CD4 T cells were used to create similar tag density plots. TSS and TTS (±4 kb)-centered plots are also shown to emphasize similarities and differences in these regions.

(B) Tag density plots for nucleosome distributions across intragenic loci stratified according to promoter CpG density in UD NCCIT (left) and CD4 T cells (right).

(C) Nucleosome distribution according to expression level in UD (left) and DF NCCIT (right) conditions.

(D) Representative MNase-seq results at two loci that show a large change in expression upon differentiation of NCCIT cells. NANOG is downregulated and this is accompanied by increased nucleosome density at the TSS. HOXB9 is upregulated and shows the opposite trend in TSS nucleosome density. The browser data enclosed by blue brackets are enlarged at the right of each panel (red boxed region), with blue ovals indicating inferred positions of nucleosomes.

(E) Analysis of relationships among DNMT occupancy, DNA-Me, and nucleosome-bound versus -free (linker) DNA. Using the method described in Kaplan et al. (2009), nucleosome-bound (red) and -free (blue) DNA was defined and the correlation between these two states and regions enriched for DNMTs and DNA-Me was calculated. Enrichments are shown as box plots, with the black bar indicating the median, the box defining the first and third quartiles of the data, and whiskers indicating the range (excluding outliers). All bound versus free (linker) comparisons for a given mark are highly significant ($p < 2 \times 10^{-16}$, Student's two-sample t test).

See also Figure S7.

based on promoter CpG density shows that HCP and ICP genes are distinct from LCP genes in NCCIT cells, with the latter displaying a peak and the former two classes showing a depletion of nucleosome density at the TSS. HCPs in CD4 T cells have the deepest NDR at the TSS, similarly to NCCIT cells, but both ICPs and LCPs show distinct nucleosome enrichments at the TSS (Figures 5B and S7C). Moderate and highly expressed gene classes were most depleted of nucleosome density at the TSS (Figure 5C). Representative nucleosome occupancy profiles at two loci displaying substantial differentiation-associated changes in expression in NCCIT cells are shown in Figure 5D, and reveal subtle changes in nucleosome density at the TSS consistent with their change in expression. CGIs were uniformly depleted of nucleosomes (Figure S7D), whereas exons were flanked by asymmetric peaks of nucleosome density (Figures S7E–S7G). Taken together, these data not only provide a comprehensive map of nucleosome positions in a human cancer cell line but also begin to reveal differences between normal and transformed cells, which may collectively reflect the impact of cell of origin and transformation state on nucleosome positioning.

We focused subsequent analyses on the relationships among nucleosome positioning, DNA methylation, and DNMT localization. Using the method described by Kaplan et al. (2009), we determined nucleosome-bound and -free (linker) DNA

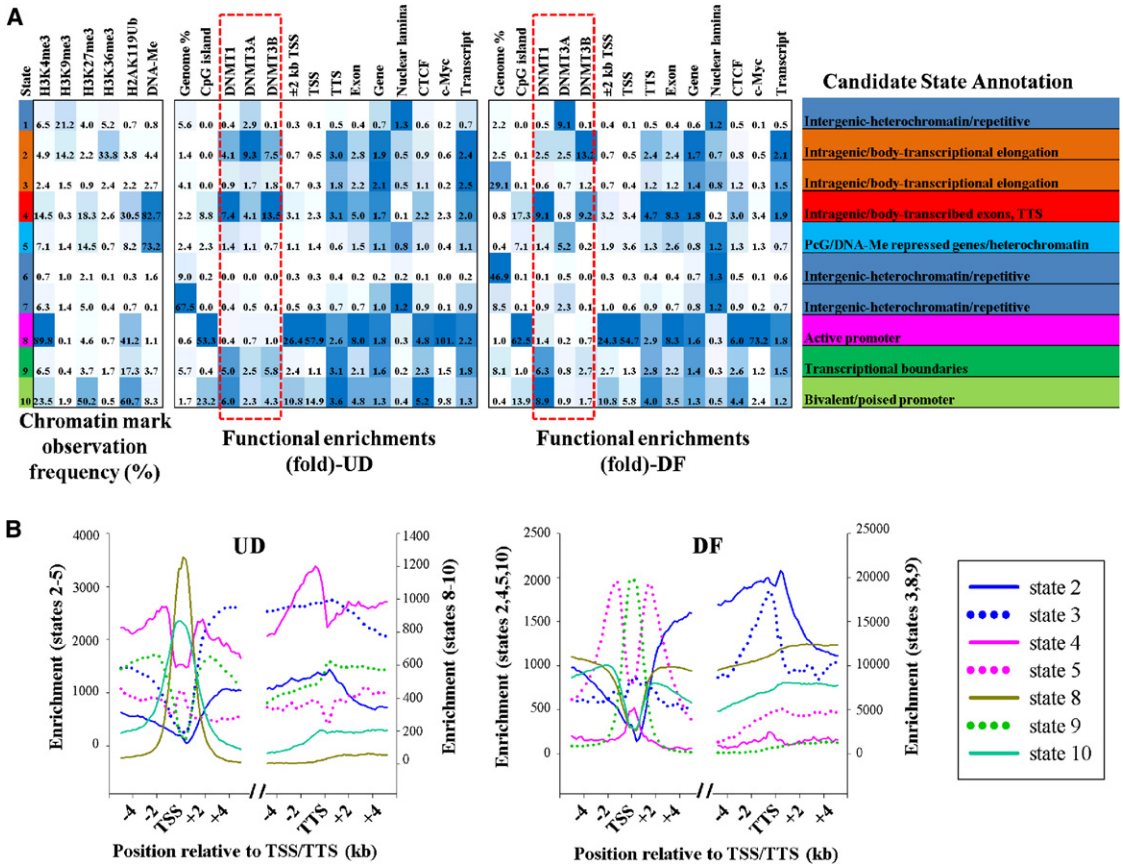

Figure 6. Discovery and Characterization of Chromatin States in NCCIT Cells

(A) A multivariate HMM was used to determine chromatin states jointly across UD and DF NCCIT cells. In the left part of the table, emission parameters that were determined de novo from recurrent combinations of epigenetic marks over the entire genome are shown. Each value represents the frequency of the epigenetic mark at genomic positions corresponding to that particular state. Functional enrichment of different features and candidate state annotations for each state in UD and DF conditions are shown on the right. Blue shading: intensity scaled by column.

(B) Comparison of enrichments for states associated with genes (states 2–5 and 8–10 only) across the TSS and the TTS regions in UD and DF cells. The break indicates the remainder of the gene body. States are graphed with different scales (right and left *y* axes) to highlight their patterns relative to other states, rather than absolute enrichments.

sequences, and then compared enrichment of DNA methylation and each DNMT in both fractions. DNA methylation was clearly enriched in the nucleosome-bound fraction. Not only did total DNA methylation within nucleosomal DNA increase after differentiation, but the difference between bound and unbound fractions also increased (Figure 5E). All DNMTs also show significantly greater enrichment in the nucleosome-bound fraction before and after differentiation (Figure 5E). Therefore, these results demonstrate that although the nucleosome is generally considered a barrier to DNA-based transactions, the DNMTs are preferentially targeted to this structure in vivo along with their mark, DNA methylation.

Defining Chromatin States Associated with DNMT Binding

To integrate all epigenetic mapping data across the two cell states and study relationships with DNMT localization, we applied a multivariate hidden Markov model (HMM; Ernst et al., 2011), which makes use of combinatorial patterns of epigenetic modifications to distinguish chromatin states. Using all chromatin modifications except nucleosome positioning, we determined the chromatin states. From these states we then selected ten that showed distinct biological enrichment (Figure 6A). These states were distinctly associated with CGIs, gene features (TSS, TTS, etc.), transcription factor binding sites (c-Myc), the multifunctional insulator protein CTCF, and regions of heterochromatin associated with the nuclear lamina (Ernst et al., 2011; Figure 6A). We plotted the enrichment of states most associated with genes (states 2–5 and 8–10) across the TSS and TTS (Figure 6B), revealing distinct patterns that further reinforce the candidate state annotations. Most of the states changed relatively little upon differentiation, except for states 3, 6, and 7 (Figure 6). DNMTs were most highly associated with states related to transcription, gene bodies, exons, and gene 3′-ends. DNMT3A also became more enriched in heterochromatin after differentiation. DNMT3B remained most highly associated with gene

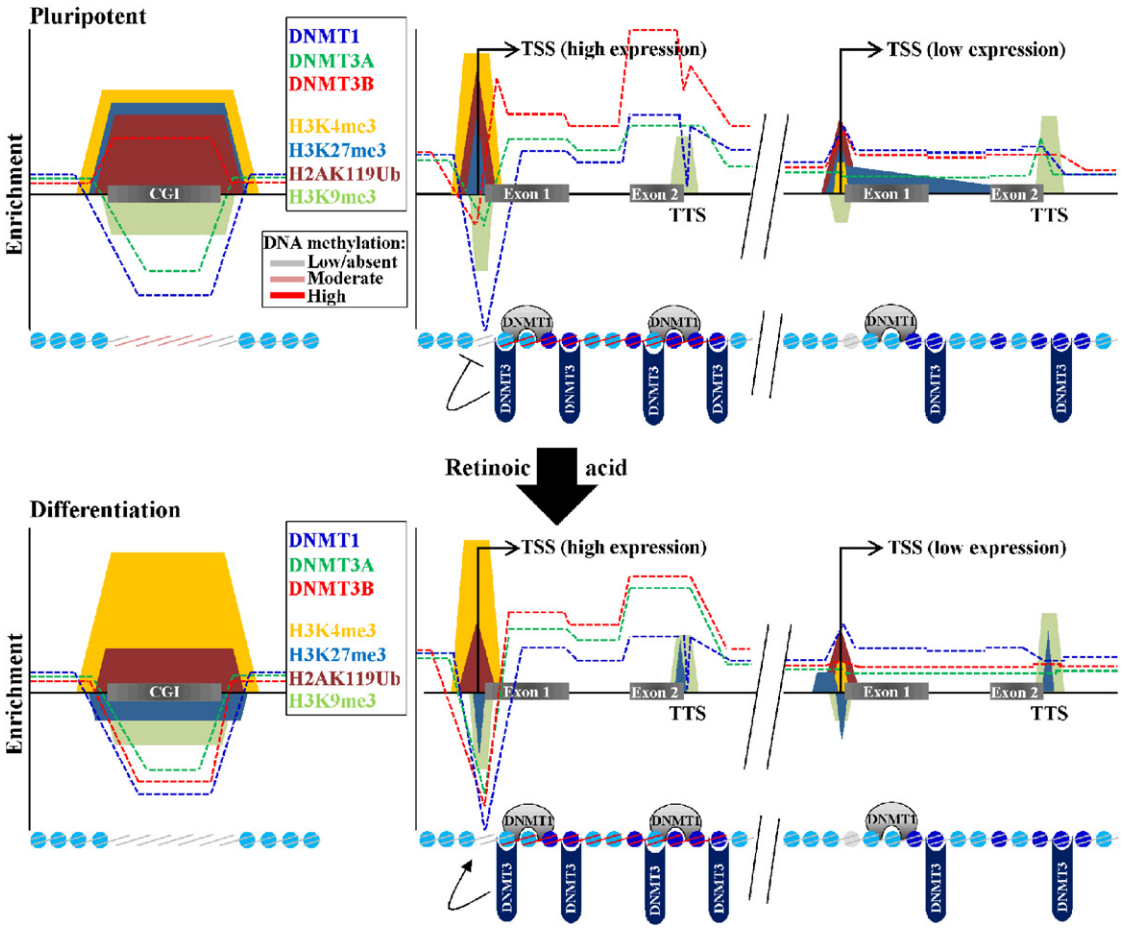

Figure 7. Summary of Key Findings and Model for Targeting DNA-Me Based on DNMT Occupancy

Summarized schematic enrichments of the indicated epigenetic marks, DNMTs, DNA-Me, and nucleosome positions in UD (top) and DF (bottom) NCCIT cells based on our findings. Shown are three features: CGIs, highly expressed loci, and lowly expressed loci. DNMT enrichments are denoted with dashed lines, histone modifications by colored shapes, nucleosomes by blue circles (darker blue indicates more highly positioned or enriched), and DNA-Me by colored lines on the nucleosomes. Note that in general, DNMTs and DNA-Me are weakly associated with CGIs, except for DNMT3B. DNMTs and DNA-Me are much more enriched at highly expressed (and CpG-rich) loci than at lowly expressed (CpG-poor) loci. The TSS is a nucleosome-free region regardless of expression level, but is less pronounced in lowly expressed genes (gray circle). After differentiation of NCCIT, loci with DNMT1 and a DNMT3 tend to acquire DNA-Me and lose the H3K27me3 PcG marker and bivalency (shown by the colored shapes on nucleosomes, based on data in Figure 3). The TSS and TTS therefore appear to set boundaries on methylation. Breakdown of these boundaries in cancer may permit aberrant methylation spreading into promoters and reduced gene-body methylation.

bodies, whereas DNMT1 was linked more to promoter regions after differentiation (Figure 6A). Our data clearly show a predominant interaction between DNMTs and DNA methylation and intragenic loci, and also that distinct classes of genes are differentially enriched for DNA methylation compared with the DNMTs (e.g., state 2 compared with state 4), further suggesting DNA-methylation-dependent and -independent roles for the DNMTs.

DISCUSSION

In this study we investigated the relationships among a critical epigenetic mark (DNA methylation), the "writers" of that mark (DNMT1, DNMT3A, and DNMT3B), histone modifications, and nucleosome structure in a cell type with relevance to both cancer and stem cell biology (summarized in Figure 7). Genome-wide mapping of five histone modifications in EC cells reveals an epigenetic landscape that displays both similarities and differences compared to hES cells. We also establish genome-wide binding profiles for all enzymatically active DNMTs and relate their binding patterns to the mark they directly mediate, DNA methylation, providing insights into gene-body methylation, possible methylation-independent functions of the DNMTs, and characteristics of loci targeted for DNA hypermethylation. Finally, we examine how the DNMTs and DNA methylation relate to the basic repeating unit of chromatin, the nucleosome, revealing an additional determinant of DNMT binding (the nucleosome itself), and distinct differences in nucleosome occupancy between normal and transformed cells. This study therefore sheds light on how DNA methylation patterns are established and maintained in the context of chromatin, which will help elucidate the epigenetic underpinnings of cancer and advance our

understanding of how pluripotency, differentiation, and induced reprogramming are regulated.

A key realization that was not fully appreciated until the DNA methylome was examined on a genome-wide scale is that this mark is not strictly linked to transcriptional repression (Figure 7). Rather, its function is dependent on genomic context: methylation in the promoter/5′-end is associated with repression, but methylation within the gene body is positively associated with expression (Maunakea et al., 2010). Although the former has been intensively studied, the latter has not, even though gene-body methylation has been recognized for some time (Jones, 1999). Our ChIP-seq data reveal strong enrichment of DNMTs in gene bodies, especially DNMT3B and DNMT1. DNMT localization patterns are similar, but not identical, to the distribution of DNA methylation. When the expression level is factored in, DNMT3B is the most highly enriched DNMT as a function of expression. Further differences between DNA methylation and DNMT binding emerged, particularly at the TSS, suggesting that DNMTs are exerting methylation-independent regulatory roles within this region. Interestingly, Dhayalan et al. (2010) reported that the PWWP domain of DNMT3A binds to H3K36me3 and enhances DNMT3A activity. It is unclear whether DNMT3B's PWWP domain has a similar preference, although an analysis of methylation in ICF syndrome cells revealed substantial gene-body hypomethylation (Aran et al., 2011). Further support for a direct link between DNMTs and transcribed loci comes from the recent report that DNMT3A and DNMT3B interact with RNA polymerase II subunits and the pol II elongation complex known as PAF1 (Rigbolt et al., 2011), which facilities transcriptional elongation and is also implicated in transcriptional initiation and messenger RNA processing (Kim et al., 2010). DNA methylation, like H3K36me, may repress spurious transcription within gene bodies that originate from repetitive elements, or regulate tissue-specific alternative promoter use or splicing. Elucidating the function of DNA methylation within the gene body and its role in disease represents a critical area of future research, which should be facilitated by the results presented here.

Data derived largely from gene-knockout studies provide strong evidence for unique and overlapping roles within the DNMT family. Our results shed significant light on this aspect of DNA methylation biology. Through integrative and comparative analyses, we define how each DNMT interacts with various genomic features, with each other, and with other epigenetic marks. These analyses show that DNMTs possess both similarities and differences in their patterns of binding. For example, DNMT3A is least enriched within intragenic loci, and DNMT3B is most enriched within gene bodies. A recent study by Choi et al. (2011) also supports the division-of-labor hypothesis. These authors showed that genes targeted for de novo methylation by ectopic DNMT3A overexpression were associated with H3K4me3, whereas those targeted by DNMT3B were associated with H3K27me3 (Choi et al., 2011). A comparison of DNMT localization with DNA methylation patterns reveals that, indeed, the majority of loci occupied by DNMTs are methylated, but also that nearly half of all methylated loci do not appear stably bound by a DNMT. The relatively small number of regions that are bound by at least one DNMT but do not overlap with methylation increases nearly 4-fold after differentiation (6-fold in gene bodies). HMM further supports a link between DNMTs and gene bodies, exons, and gene 3′-ends. The HMM shows that some states have high DNA methylation and DNMT representation (e.g., state 4), while other states do not show equal representation (e.g., states 2, 9, and 10). Indeed, there is growing evidence that DNMTs serve non-DNA-methylation-related functions in gene regulation (Espada, 2012). Alternatively, loci that are occupied by DNMT but are not methylated may be primed for later de novo methylation events, a notion that is supported by our comparison of the epigenomes of NCCIT cells in pluripotent and DF states and with hES cells (Figure 7). It is also possible that certain biases inherent to all methylated DNA enrichment-NGS methods and challenges associated with mapping reads within repetitive loci contribute to our finding of incomplete overlap between methylated regions, DNMT-bound loci, and heterochromatic repetitive regions. Despite these potential issues, however, our data reveal the existence of consistent links between DNMTs and transcribed loci, suggesting that DNMTs and DNA methylation have important but as yet unrecognized functions within intragenic regions and are not exclusively associated with heterochromatin or transcriptional repression.

In a recent study, ~20% of the genome exhibited a significant nucleosome positioning preference, suggesting that the majority of nucleosomes are not subject to rigorous positioning, and that most promoters have an NDR (Valouev et al., 2011). Our data confirm the presence of a genome-wide NDR flanking the TSS, and we observed further depletion as a function of expression. By comparing NCCIT nucleosome positions with published nucleosome positioning data derived from CD4 T cells, we identified several differences. The TTS region in particular showed opposing trends in nucleosome density, and it will be of interest to determine the contribution of cellular transformation to this and other differences as more nucleosome profiles are generated. Our results show that DNMTs, like methylation (Chodavarapu et al., 2010), are more highly enriched in nucleosome-bound DNA compared with linker DNA, and methylation becomes further enriched in nucleosomes after differentiation (Figure 7). Exactly how DNMTs are targeted to or access CpG sites within nucleosomal DNA is unknown. Chromatin remodelers that interact with DNMTs, such as SNF2H (Geiman et al., 2004), could facilitate this process. Alternatively, mechanisms may exist to couple DNA methylation to nucleosome reassembly following DNA replication, perhaps involving interactions among DNMTs, histone modifiers, and CAF1 (Sarraf and Stancheva, 2004).

Our results and those of others (e.g., Barski et al., 2007; Ernst et al., 2011) reveal distinct boundaries at the TSS and TTS at the chromatin level. The data presented here show clear partitioning of DNMTs between promoters and gene bodies as well. It is intriguing to speculate that a defect in the TSS and/or TTS chromatin boundaries during tumorigenesis, or perhaps weaker boundaries in stem cells due to their more plastic nature, might permit DNA methylation/DNMTs to encroach into the promoter, or permit activation of methylation at DNMT-bound but normally unmethylated loci (Figure 7). Some of these methylation events may be part of normal cell-type-specific epigenome changes associated with differentiation, leading defects in this process to cause methylation destined for one cell type to be aberrantly imposed on another (Irizarry et al., 2009). A boundary defect

could account for gene-body hypomethylation as well, due either to a breakdown in the targeting mechanisms or simple dilution of DNMTs. Several of the HMM states enriched for DNMTs, such as states 2 and 4, peak just inside the TSS and TTS. By examining chromatin characteristics in two different systems, we were able to show that the presence of at least two DNMTs, along with other epigenetic marks such as those mediated by PcG complexes and H3K9me3, mark loci targeted for DNA hypermethylation in settings relevant to both adult cancers and differentiation programs (Figure 7). Interestingly, both the maintenance DNMT and at least one de novo DNMT are enriched at methylated loci prior to the increase in methylation. DNMT3B represents a particularly strong candidate initiator of cancer-specific methylation changes because it is the only DNMT that is enriched at CGIs and is highly overrepresented at HCP genes that increase in methylation. Mapping DNMTs and chromatin marks in other cell types will be critical to refine models of methylation targeting, and has great potential to uncover previously unknown mechanisms and corrective measures for the methylation defects that typify cancer cells.

EXPERIMENTAL PROCEDURES

Cell Culture, Treatments, RNA Extraction, and qPCR

NCCIT (from ATCC) and WA09 (H9, from WiCell) cells were grown as described in Extended Experimental Procedures. All-*trans* RA (Sigma) was dissolved in DMSO. Cells were treated with 10 μM RA or an equivalent volume of DMSO and collected at day 0, 1, 4, 7, 14, or 21. Total RNA preparation and RT-qPCR were performed as described in Gopalakrishnan et al. (2009) and Extended Experimental Procedures.

Expression Microarrays

Gene expression profiling was performed using Affymetrix Human Gene 1.0 ST arrays. All samples were run in triplicate at the GHSU Cancer Center Genomics Core Facility and analyzed as described previously (Jin et al., 2009) and in Extended Experimental Procedures.

ChIP and Western Blotting

ChIP and western blotting were performed essentially as described previously (Jin et al., 2009). Detailed conditions and a list of all antibodies used are provided in Extended Experimental Procedures.

MBD Methylated DNA Enrichment

Enrichment of CpG methylated DNA for NGS was based on the method originally described in Serre et al. (2010) with the modifications detailed in Extended Experimental Procedures.

Nucleosome Preparation

Genome-wide nucleosome positioning experiments (MNase-seq) were performed essentially as described in Schones et al. (2008). Additional details are provided in Extended Experimental Procedures.

Library Preparation and NGS

DNA fragments isolated by ChIP, MBD enrichment, or MNase digestion were used to create single-end sequencing libraries using the Genomic DNA Sample Prep Kit (Illumina) with only minor modifications, as described in Extended Experimental Procedures. Libraries were quality controlled on a Bioanalyzer 2100 high-sensitivity DNA chip and sequenced on an Illumina GAIIX (72 bp read) or a HiSeq2000 (100 bp read) at the Tufts University Genomics Core Facility or at BGI Americas.

Data Analysis and HMM Modeling

Raw sequence data were processed using the default settings for base calling to ensure read quality, and reads containing more than two unidentifiable bases were discarded, producing usable sequence data in standard FASTQ format. Reads were mapped against the human reference genome (hg19) using Novoalign software (http://www.novocraft.com) to produce alignment files. Alignment files were further analyzed by MACS1.4 (http://liulab.dfci.harvard.edu/MACS) for peak calling. Further processing and analysis of sequence data derived from all ChIP-seq, MBD-seq, and MNase-seq experiments are described in the Extended Experimental Procedures. HMM using the histone modification and DNA methylation data was performed as described previously (Ernst et al., 2011) and in the Extended Experimental Procedures.

ACCESSION NUMBERS

Sequencing and expression microarray data have been deposited into the NCBI Gene Expression Omnibus database under accession number GSE38938.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2012.10.017.

LICENSING INFORMATION

ACKNOWLEDGMENTS

We thank Kip Bodi and the Tufts University Genomics Core Facility for assistance with Illumina sequencing, and Keji Zhao (National Heart, Lung, and Blood Institute) for providing advice on nucleosome positioning experiments. This work was supported by the National Institutes of Health (grants R01CA114229, R01CA116028, and R01AA019976 to K.D.R.; RC1HG005334 to M.K.; and P01GM08535403 to S.D. and K.D.R. [pilot fund]) and the National Science Foundation (postdoctoral fellowship 0905968 to J.E.). K.D.R. is a Georgia Cancer Coalition Distinguished Cancer Scholar.

Received: April 12, 2012
Revised: June 30, 2012
Accepted: October 22, 2012
Published: November 21, 2012

REFERENCES

Adamo, A., Sesé, B., Boue, S., Castaño, J., Paramonov, I., Barrero, M.J., and Izpisua Belmonte, J.C. (2011). LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. Nat. Cell Biol. *13*, 652–659.

Aran, D., Toperoff, G., Rosenberg, M., and Hellman, A. (2011). Replication timing-related and gene body-specific methylation of active human genes. Hum. Mol. Genet. *20*, 670–680.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. Nat. Rev. Cancer *11*, 726–734.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407–412.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.,

et al.; ENCODE Project Consortium; NISC Comparative Sequencing Program; Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.-Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J., et al. (2010). Relationship between nucleosome positioning and DNA methylation. Nature *466*, 388–392.

Choi, S.H., Heo, K., Byun, H.-M., An, W., Lu, W., and Yang, A.S. (2011). Identification of preferential target sites for human DNA methyltransferases. Nucleic Acids Res. *39*, 104–118.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J. Biol. Chem. *285*, 26114–26120.

Ehrlich, M., Jackson, K., and Weemaes, C. (2006). Immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF). Orphanet J. Rare Dis. *1*, 2.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.

Espada, J. (2012). Non-catalytic functions of DNMT1. Epigenetics *7*, 115–118.

Gao, Q., Steine, E.J., Barrasa, M.I., Hockemeyer, D., Pawlak, M., Fu, D., Reddy, S., Bell, G.W., and Jaenisch, R. (2011). Deletion of the de novo DNA methyltransferase Dnmt3a promotes lung tumor progression. Proc. Natl. Acad. Sci. USA *108*, 18061–18066.

Geiman, T.M., Sankpal, U.T., Robertson, A.K., Chen, Y., Mazumdar, M., Heale, J.T., Schmiesing, J.A., Kim, W., Yokomori, K., Zhao, Y., and Robertson, K.D. (2004). Isolation and characterization of a novel DNA methyltransferase complex linking DNMT3B with components of the mitotic chromosome condensation machinery. Nucleic Acids Res. *32*, 2716–2729.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. *74*, 481–514.

Gopalakrishnan, S., Sullivan, B.A., Trazzi, S., Della Valle, G., and Robertson, K.D. (2009). DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. Hum. Mol. Genet. *18*, 3178–3193.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat. Genet. *41*, 178–186.

Jin, B., Yao, B., Li, J.-L., Fields, C.R., Delmas, A.L., Liu, C., and Robertson, K.D. (2009). DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer. Cancer Res. *69*, 7412–7421.

Jones, P.A. (1999). The DNA methylation paradox. Trends Genet. *15*, 34–37.

Kallin, E.M., Cao, R., Jothi, R., Xia, K., Cui, K., Zhao, K., and Zhang, Y. (2009). Genome-wide uH2A localization analysis highlights Bmi1-dependent deposition of the mark at repressed genes. PLoS Genet. *5*, e1000506.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., and Segal, E. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362–366.

Kim, J., Guermah, M., and Roeder, R.G. (2010). The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. Cell *140*, 491–503.

Laird, P.W., Jackson-Grusby, L., Fazeli, A., Dickinson, S.L., Jung, W.E., Li, E., Weinberg, R.A., and Jaenisch, R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. Cell *81*, 197–205.

Linhart, H.G., Lin, H., Yamada, Y., Moran, E., Steine, E.J., Gokhale, S., Lo, G., Cantu, E., Ehrich, M., He, T., et al. (2007). Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. Genes Dev. *21*, 3110–3122.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature *466*, 253–257.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W., 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.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases *Dnmt3a* and *Dnmt3b* are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Rigbolt, K.T.G., Prokhorova, T.A., Akimov, V., Henningsen, J., Johansen, P.T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J.V., and Blagoev, B. (2011). System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci. Signal. *4*, rs3.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol. Cell *15*, 595–605.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E., 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.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887–898.

Serre, D., Lee, B.H., and Ting, A.H. (2010). MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res. *38*, 391–399.

Sperger, J.M., Chen, X., Draper, J.S., Antosiewicz, J.E., Chon, C.H., Jones, S.B., Brooks, J.D., Andrews, P.W., Brown, P.O., and Thomson, J.A. (2003). Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proc. Natl. Acad. Sci. USA *100*, 13350–13355.

Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. Nature *474*, 516–520.

Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871–874.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903.