Title: DNMT inhibitors increase methylation at subset of CpGs in colon, bladder, lymphoma, breast, and ovarian, cancer genome

Running title: Decitabine/azacytidine increases DNA methylation

Anil K Giri[1], Tero Aittokallio[1,2]

[1]Institute for Molecular Medicine Finland, FIMM, University of Helsinki, Helsinki, Finland.

[2]Department of Mathematics and Statistics, University of Turku, Turku, Finland.

Correspondence to

Dr. Anil K Giri

Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland.

Email: anil.kumar@helsinki.fi

Financial disclosure: This work was funded by the Academy of Finland (grants 269862, 292611, 310507 and 313267), Cancer Society of Finland, and the Sigrid Juselius Foundation.

Ethical disclosure: This study is an independent analysis of existing data available in the public domain and does not involve any animal or human samples that have been collected by the authors themselves.

Author contribution: AKG conceptualized, analyzed the data and wrote the manuscript. TA critically revised and edited the manuscript. The authors report no conflict of interest.

Abstract

Background: DNA methyltransferase inhibitors (DNMTi) decitabine and azacytidine are approved therapies for acute myeloid leukemia and myelodysplastic syndrome. Identification of CpGs violating demethylaion due to DNMTi treatment may help to understand their resistance mechanisms.

Materials and Methods: To identify such CpGs, we analysed publicly available 450K methylation data of multiple cancer type cell lines.

Results: We identified 637 CpGs corresponding to genes enriched for p53 and olfactory receptor pathways with a transient increase in methylation (median $\Delta\beta = 0.12$) after decitabine treatment in HCT116 cells. Azacytidine treatment also increased methylation of identified CpGs in 9 colon, 9 ovarian, 3 breast, and 1 lymphoma cancer cell lines.

Conclusion: DNMTi treatment increases methylation of subset of CpGs in cancer genome.

Keywords

Decitabine, azacytidine, methylation, colon cancer, RORA, HCT116, olfactory receptor pathways, alternative splicing

Introduction

DNA methyltransferase inhibitors (DNMTi) are widely used as chemical tools for hypomethylating the genome in order to understand the role of DNA methylations in X-chromosome inactivation, DNA imprinting and transcriptional regulation of several disease-related genes [1-4]. Further, DNMTi agents, decitabine along with its analog azacytidine, have been approved by United States Food and Drug Administration (US FDA), and they currently remain as the sole treatment option for specific sub-groups of acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) patients [5-6]. Since DNA methylation-induced silencing of tumor suppressor genes, such as P53, at promoter region is a primary event in many cancers and these methylations can be reversed by DNMTi as therapy, both of these drugs are also being tested as a treatment option for breast, lung, colon and other cancers. Decitabine treatment causes global hypomethylation of the genome by intercalating itself in the DNA during replication and halting the DNA methylation transferases (DNMTs) actions [5-6]. Hypomethylation of the genome leads to re-expression of several genes, including multiple tumor suppressor and inhibition of oncogenes, thereby contributing to apoptosis of cancer cells through multiple ways such as DNA damage response pathway, p53 signaling pathways, cytotoxicity, etc [6,7].

However, there are sporadic reports where treatment with DNMTi has led to an increased expression level of DNA methylating enzymes hence DNA methylation in specific cells [8-11]. For example, Kastl *et al*. reported an increase in the mRNA level of DNMT1, DNMT3a and DNMT3b genes in docetaxel-resistant MCF7 cells as compared to drug sensitive cells when treated

with decitabine [8]. Surprisingly, a recent study showed that decitabine treatment can cause an increase in 5-hydroxymethylcytosine, an oxidation product of methylated cytosine, in DNA of human leukemic cells [9]. Further, an analog of decitabine, azacytidine treatment, was reported to induce DNA methylation in transgenes of Chinese hamster cell in the process of silencing foreign genes in the human genomes [10]. This piece of evidence hints that treatment with azacytidine can induce DNA methylation at certain locations in the genome that may have non-human origins such as retrotransposons and other genes with viral origin [10,11]. Available piece of literature also suggests that DNMTi treatment causes hypomethylation nearly at 99% of methylated locations in the genome [12], suggesting that there should also be loci where DNMTi treatment can increase the methylation level or has no effect on methylation, instead of the regular role of hypomethylation. However, we are currently lacking the information of the genomic location, function, origin, and fate of those CpGs in the cancer genome that can resist the DNA demethylation.

In the present work, we aim to systematically investigate the extent, location and role of CpGs with increased methylation in response to DNMTi treatment. Identification of such loci and their related genomic features will not only help to understand the reasons behind the failure of the DNMTi treatment in demethylating cancer-related genes but it may also reveal novel molecular mechanism behind efficacy, side effects, and resistance towards DNMTi treatment in various cancer types. We selected HCT116 cell line as our primary disease model to discover these CpGs as it shows the silencing of various tumor suppressor genes due to hypermethylation as seen in the case of colon cancer tissue [13]. Further, HCT116 cell line has been frequently utilized to study DNA methylation and its role in regulating gene expression in colon cancer [14]. To investigate how general these findings are, we tested the increase in methylation after DNMTi treatment identified in HCT116 cells also in other lymphoma, colon, ovarian, and breast cancer cells. Further, we explored the relationship between methylation status of the identified loci and expression status of genes in colon adenocarcinoma cancer using patient tumor data from The Cancer Genome Atlas

(TCGA) project. Our work lays foundation for the search of rare events of hypermethylation due to DNMTi treatment contrary to their classic role of DNA hypomethylation in the cancer genome.

Methodology

Processing of methylation data

To identify CpGs with increased methylation after decitabine treatment we analyzed the DNA methylation (Illumina 450K platform, GSE51810) and gene expression data (Illumina HumanHT-12_V4_0_R1 platform, GSE51810) from the study by Yang *et al*. [15] for HCT116 colon cell lines treated with decitabine (0.3 mM) for 72 hours. Cells were maintained in McCoy's 5A medium, supplemented with 10% fetal bovine serum along with 1% penicillin/streptomycin after drug treatment, and followed through 5, 14, 24, 42, and 68 days. The increase in DNA methylation in HCT116 cells were validated using methylation data from the study by Han *et al* [16] (Illumina 450K, GSE41525), where HCT116 and T24 (bladder cancer) cell lines were treated with 0.3 µM and 1 µM of decitabine, respectively for 24 hours and Illumina 450K assay was performed for both untreated and decitabine treated cells. We also tested the increase in DNA methylation of identified CpGs using DMSO (as mock) and decitabine-treated MCF7 cells in data generated by Leadem *et al* (Illumina 450K platform, GSE97483) [17]. These cells were cultured in Minimum Essential Medium (MEM) with 10% fetal bovine serum and treated with 0.06 µM of decitabine for 72 hours.

We also extended our findings discovered in case of decitabine in another DNMTi inhibitor, azacytidine, by analyzing DNA methylation data (Illumina 450K, GSE45707) for untreated and azacytidine-treated (5mM for 72 hours) lymphoma cancer U937 cell line. We further analysed additional methylation data for 26 breast cancer cell lines (MDA231,SKBR3, HCC38, ZR7530, HCC1937, CAMA1, MDA415, HCC1500, BT474, EFM192A, MDA175, MDA468, MDA361, HCC1954, BT20, ZR751, HCC1569, EFM19, T47D, MDA453, MCF7, HCC1187, HCC1419, EFM192A, MDA436, SUM149, and SUM159), 12 colorectal cancer cell lines (SW48, HCT116, HT29, RKO, SW480, Colo320, Colo205, SW620, SNUC-1,CACO-2, SK-CO1, and Colo201), and 13 ovarian cancer cell lines (TykNu, CAOV3, OAW28, OV2008, ES2, EF27, Kuramochi, OVKATE , Hey, A2780, ES2, OVCAR3, OVCAR5, and SKOV3) measured

after mock treatment and 0.5 μM azacytidine treatment for 72 hours (Illumina 450K platform, GSE57342). The cells have been cultured and maintained under recommended conditions for each cell line [18].

To investigate the alteration in methylation status of identified probes in cancerous tissue and their role in gene expression regulation, TCGA level 3 HumanMethylation 450K data and normalized RNA-seq gene expression profiles for colon adenocarcinoma (COAD) samples were downloaded using the FireBrowse tool (http://gdac.broadinstitute.org/).

Methylation status at a CpG site was measured as beta value (β) which is the ratio of the methylated probe intensity and the overall intensity (sum of methylated and unmethylated probe intensities designed for a particular CpG in 450K beadchip). It ranges from 0 to 1, indicating no methylation (β=0) to complete methylation of the CpGs (β=1). We performed appropriate quality control of the published data before their downstream analysis. We removed all the CpGs with missing values and a tendency of cross-hybridization as specified in the supplementary file of Chen *et al* [19]. To remove any possible bias due to design differences in the type of probes (the type I and type II probes) present in the Illumina 450K platform, we performed BMIQ normalization [20] to the DNA methylation data for TCGA samples before correlation and differential methylation analysis. All other data processing was done using local inbuilt commands in R as described previously [21].

Identification of probes with increased methylation in HCT116 cell line

We calculated the difference in methylation level of CpGs before and after treatment with decitabine in HCT116 cell line at day 5 in data from Yang *et al* (GSE51810). CpG that showed an increase in β-value of greater than or equal to 0.10 ($\Delta\beta \geq 0.10$) between untreated control and decitabine treated HCT116 cells after 5 days were identified as CpGs with increased methylations.

Gene expression data analysis for HCT116 cell line and colon adenocarcinoma tumors from TCGA

Expression analysis was carried out using the inbuilt commands in R. The data were $\log_2$-transformed and normalized using Robust Spline Normalization (RSN) using the lumi package in R

[22]. Pearson correlation between methylation and gene expression across different time point in HCT116 cell line was assessed using the cor.test function of R.

Gene expression data from TCGA samples were normalized using voom function in the limma package [23], and these data were Z-transformed before the differential and correlation analyses. Wilcoxon non-parametric test was used to identify differentially expressed genes in TCGA adenocarcinoma samples (FDR<0.05). Only those adenocarcinoma samples that had both DNA methylation and gene expression information were used for the correlation analyses.

Gene annotation and pathway enrichment analysis

Identified CpGs were annotated for their location in the genome based on annotation file provided by Illumina (ftp://ussd-ftp.illumina.com/downloads/ProductFiles/HumanMethylation450/HumanMethylation450_15017482_v1-2.csv). Gene ontology and pathway enrichment analysis of genes corresponding to CpGs with increased methylation were done using GeneCodis [24]. Statistical enrichment was assessed by FDR corrected p-values from hypergeometric test for separate ontology terms and pathways. GENEMANIA [25] was used to construct and visualize the interaction network between genes and the transcription factor regulating them. Key term enrichment analysis for the genes corresponding to CpGs was done using DAVID [26].

Results

DNMTi treatment causes induction of methylation in a small portion of the CpGs

After quality control, we analyzed 369,886 CpGs across the genome of HCT116 cells from Yang *et al*. [15], and identified hypermethylation ($\Delta\beta \geq 0.10$) of 638 unique CpGs (0.02% of the total analyzed CpGs) in 393 unique genes after 5 days of decitabine treatment, as compared to untreated cells (Figure 1A). Most of them were hypomethylated in the untreated state (median $\beta = 0.18$), and after decitabine treatment, a median increase of 0.12 ($\Delta\beta = 0.12$) in methylation level was observed for these sites. The detailed list of the identified CpGs is provided in Supplementary Table 1.

Analysis of another methylation data for HCT116 cell line from the Han *et al* study [16] validated our finding, as we found a corresponding increase in methylation level (median

increase in β values 0.09) after decitabine treatment (0.3 µM for 24 hours) at 583 (91%) differentially methylated CpGs that are common between the two studies (Figure 1B). These results indicate that the increase in DNA methylation of most of the identified sites starts as early as 24 hours after the DNMTi treatment and lasts up to at least day 5. The findings are also robust to common technical and processing variability across two laboratory conditions.

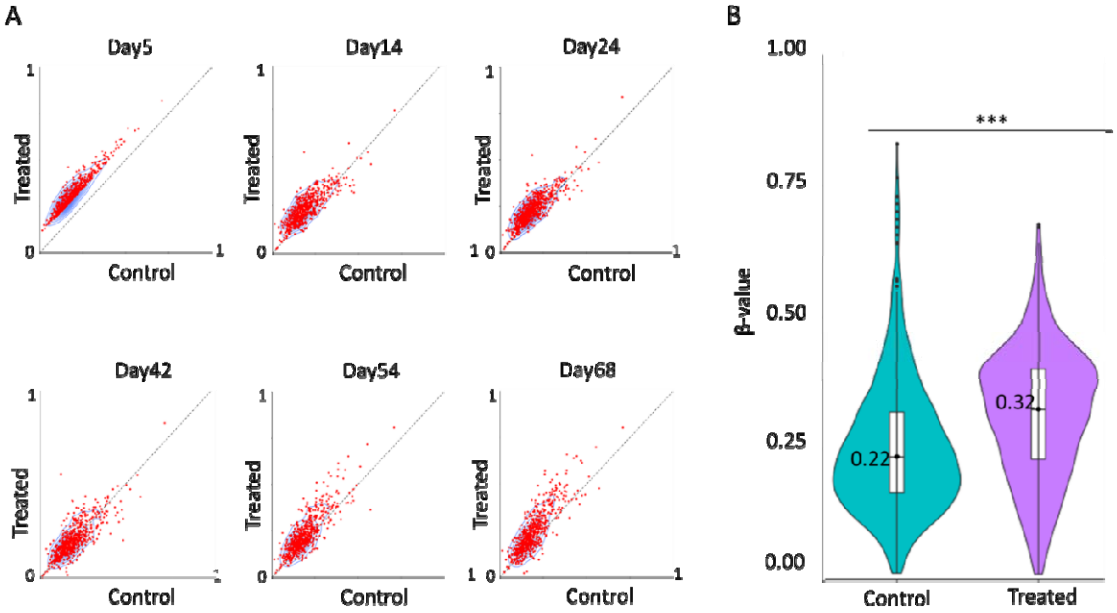


Figure 1: decitabine treatment increases DNA methylation levels of a subset of CpGs. (A) Scatter plots showing DNA methylation patterns of 638 differentially methylated CpGs between untreated control cells and decitabine treated cells at various time points in Yang *et al*. study [15]. *The x-axis indicates the DNA methylation level of probes in the untreated control, and the y-axis in decitabine treated cells. (B) Violin plot showing the median methylation level (horizontal line) and distribution patterns (density and IQR) of the identified 583 CpGs in untreated and decitabine treated HCT116 cells after 24 hours in* Han *et al* study [16]. *The statistical significance was assessed using the non-parametric Wilcoxon test. ***P<0.0005.*

Increase in methylation of identified CpGs is cancer type and tissue-specific

To test the effect of decitabine treatment on identified differentially methylated CpGs in other cancer cell lines, we re-analyzed publicly available data for decitabine-treated bladder cancer T24 cells. An increase in median DNA methylation levels ($\Delta\beta$ =0.14) at 616 (97%) common CpGs was observed after drug treatment (1 µM of decitabine for 24 hours) in T24 cells (Figure 2A). We further analyzed methylation level of identified loci for breast cancer cell line (MCF7), where cells have been treated with 0.06 µM of decitabine for 72 hours, but did not observe decrease in median methylation level ($\Delta\beta$ = -0.01) of 590 common differentially methylated CpGs (93%) in response to decitabine treatment (Figure 2B). These results indicate that methylation levels of the identified probes either increase after decitabine treatment in multiple cancer types or remain similar which is in contrast to the general effect of decitabine over CpGs methylation as we observed a significant decrease in the median methylation ($\Delta\beta$ = -0.14 for T24 and $\Delta\beta$ = -0.22 for MCF7) level of other CpGs present in the 450K beadchip in both of the cell lines (Supplementary Figure 1). Further, the result suggests that the degree of the change in methylation of these probes is highly cancer-specific.

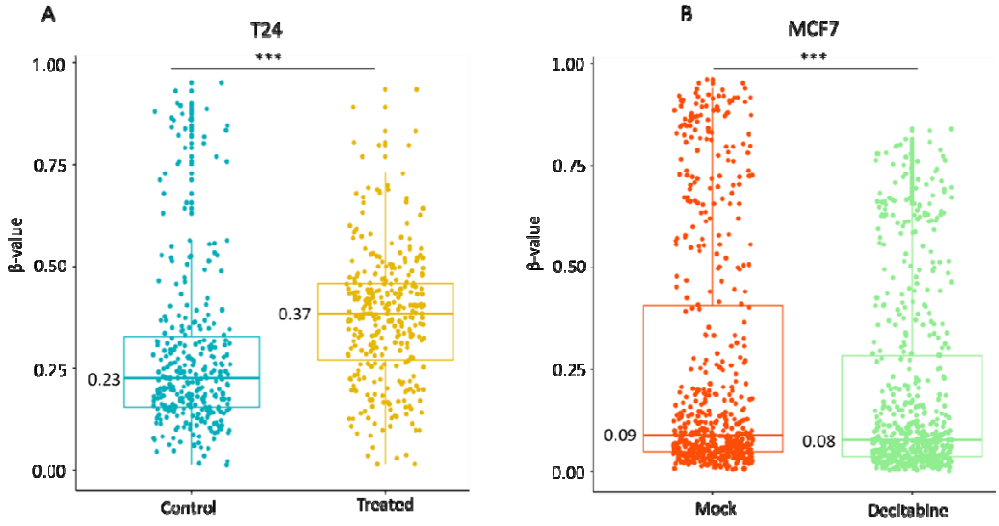


*Figure 2: Increase in methylation of identified CpGs is tissue specific. (A) Methylation level of 616 identified probes in untreated and decitabine-treated bladder cancer T24 cell line after 24 hours of drug treatment. (B) Methylation level of 590 identified probes in mock (DMSO) treated and decitabine-treated breast cancer MCF7 cell line after 72 hours of drug treatment. The statistical significance was assessed using the non-parametric Wilcoxon test. ***P<0.0005*

DNMTi treatment also increased methylation level of identified CpGs in colon, ovarian and breast cancer cells

To study the question whether the increase in methylation of the identified sites is decitabine-specific or whether also another DNMT inhibitor shows similar changes, we tested the methylation induction behavior of azacytidine, another FDA approved DNMT inhibitor, in multiple cancer cell lines. An analysis of 450K data from 52 cell lines revealed that azacytidine also increased methylation of identified CpGs in 9 out of 13 (69%) ovarian cancer cell lines(median $\Delta\beta$ >0.03), 3 out of 26 (11.5%) breast cancer cell lines(median $\Delta\beta$ >0.01), and in 9 out of 12 (75%) colon cancer cell lines (median $\Delta\beta$ >0.02), and 1 lymphoma cell line (U937, median $\Delta\beta$ >0.06) as shown in Figure 3A,B,C and D. Our analysis revealed that the increase in methylation level of identified loci in response to azacytidine treatment is not universal across all the cancer cell types, rather azacytidine treatment also causes an increase in median DNA methylation of identified sites in a tissue-specific manner.

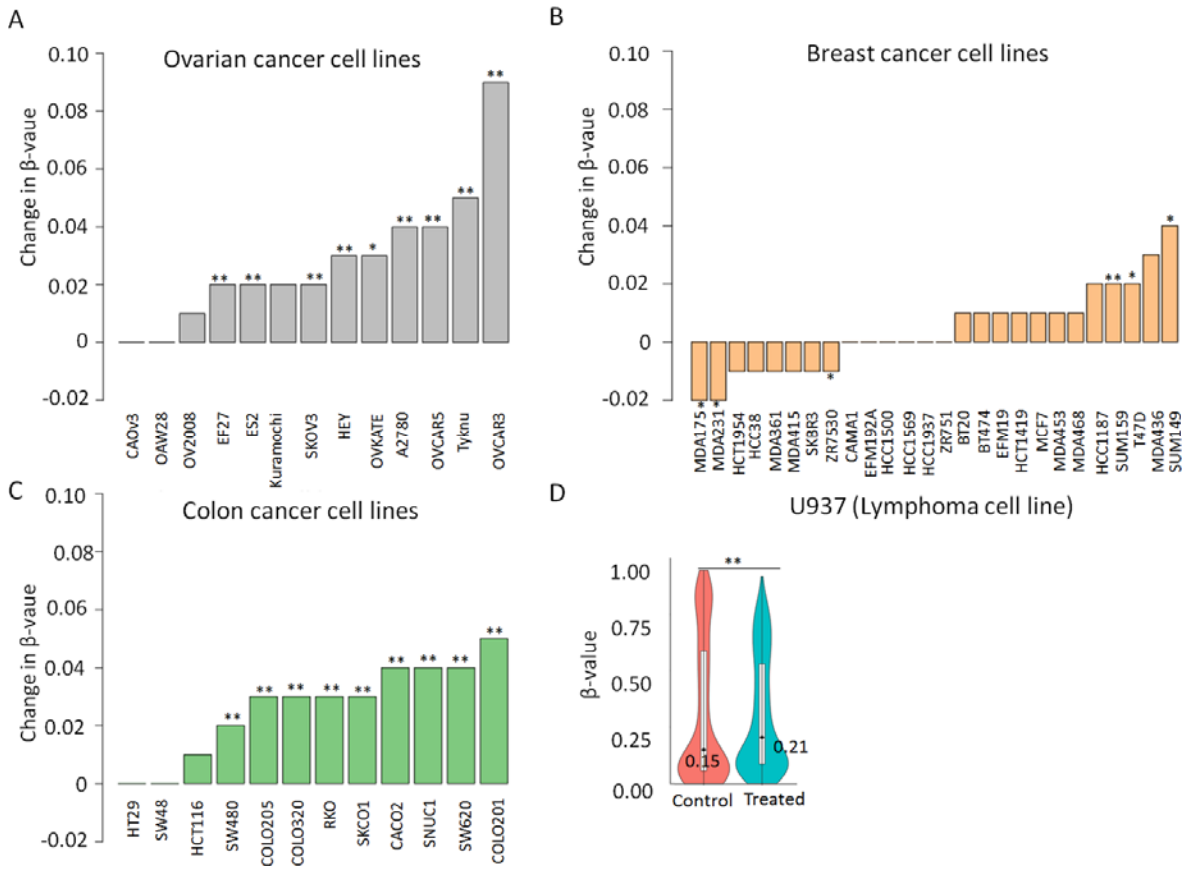


*Figure 3: Azacytidine treatment increases methylation of identified sites in a subset of cell lines. Change in median methylation level of identified CpGs in (A) 13 ovarian cancer cell lines (B) 26 breast cancer cell lines (C) and 12 colon cancer cell lines has been shown as barplot. The bar plot represents the difference in the median methylation level of identified CpGs between cells treated with 0. 5 µM azacytidine (test group) or carboplatin (mock group) after 72 hours. (D) Violin plot showing the distribution of methylation level of identified probes in untreated (control) and treated cells (*5mM for azacytidine for 72 hours) *in U937 lymphoma cell lines. The statistical significance was assessed using the non-parametric Wilcoxon test. *P<0.05, **P<0.005*

Genes corresponding to the identified CpGs show differential expression in cancerous tissue

One of the common mechanism how DNA methylation affects the biological processes is by modulating the expression of the nearby genes. To test the functional role of the increase in methylation of the identified differentially methylated CpGs, we explored the correlation between DNA methylation and expression levels of the corresponding genes using data from HCT116 cell line after decitabine treatment at multiple time-points (0, 5, 14, 24, and 42 days). A strong correlation ($|r| \geq 0.80$) between the gene expression and DNA methylation profiles was observed at 26% (N = 166) loci in HCT116 cells, out of them, 48 CpGs corresponding to 43 genes were significant ($P<0.05$). We observed a highly-significant correlation between the gene body CpGs (cg08099431) methylation in RORA gene and its expression level in HCT116 cells ($r = 0.99$, FDR = 0.01) in HCT116 (Supplementary Figure 2).The correlation plot for highly correlated CpGs ($|r| \geq 0.80$) falling in promoter and gene body regions in HCT116 cell line is shown separately in Figure 4A and 4D, respectively.

To investigate the clinical relevance of increase in methylation of identified CpGs, we analyzed the methylation level of the 109 common CpGs showing strong expression-methylation correlation ($|r| \geq 0.80$) in HCT116 cell line in TCGA colon adenocarcinoma samples. This analysis revealed that 43% (47 out of 109) of the correlated CpGs were also differentially methylated between healthy and cancerous colon tissues in the TCGA data (FDR <0.05, Figure 4B and E, Supplementary Table 2). Differential expression analysis revealed that 77% of the corresponding genes (N=83 out of 112 genes for which data was available in TCGA) were also differentially expressed between the colon and normal tissues (FDR <0.05, Figure 4C, Supplementary Table 3). The differential expression of genes corresponding to the identified CpGs in cancerous tissues indicates that the increase in their DNA methylation is pathological in colon cancer.

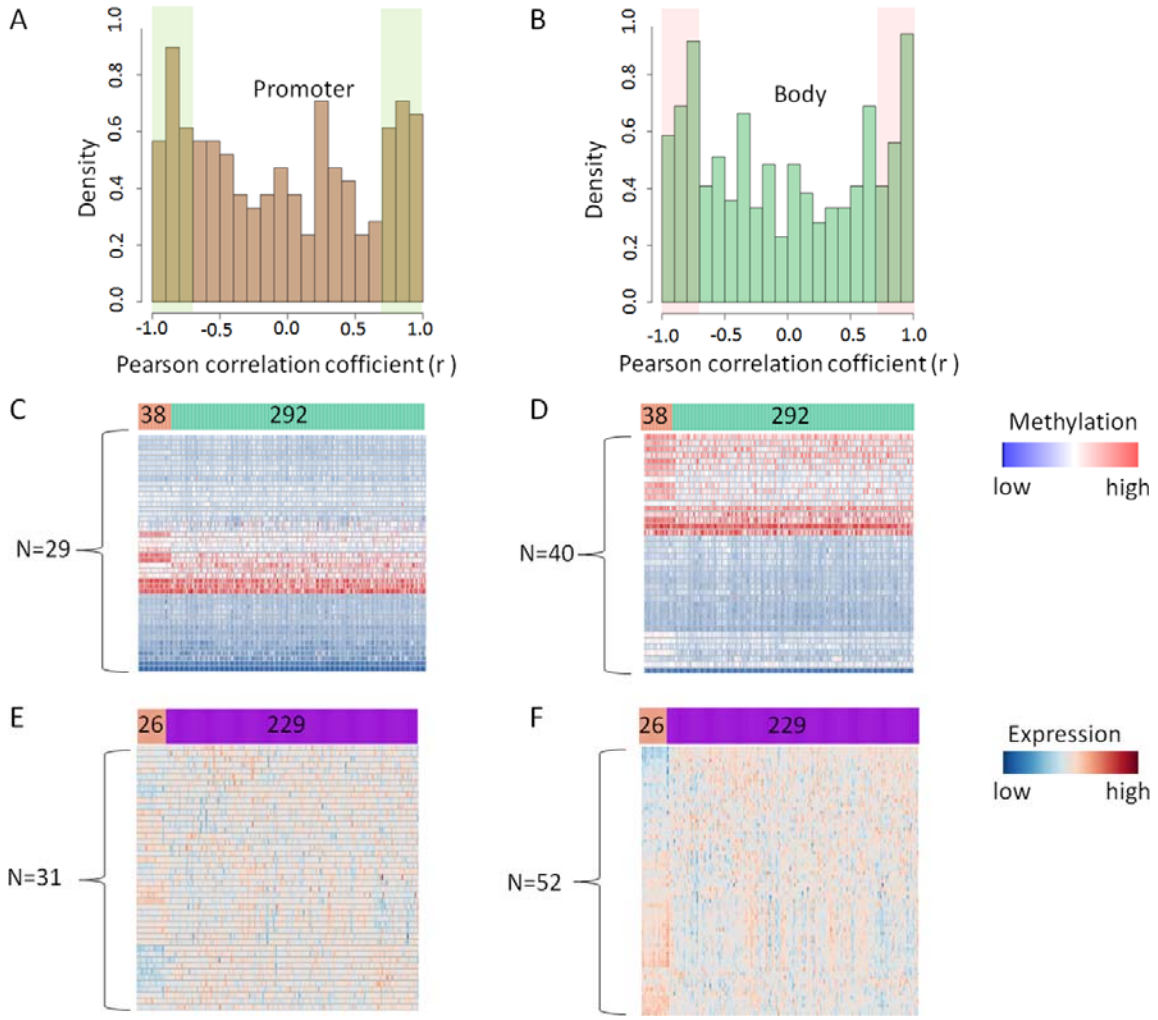


Figure 4: DNA methylation and gene expression analysis of identified CpGs and showing a strong correlation (|r|≥0.80) between expression and methylation in HCTT116 cell line in TCGA colon adenocarcinoma samples.

(A) Distribution of correlation between DNA methylation and expression at the promoter region (left panel) and gene body (right panel) in HCT116 cell line. The x-axis denotes the Pearson correlation coefficient between gene expression and methylation in HCT116 cell lines and the y-axis the kernel density (B) Heatmap showing methylation level of the identified CpGs in TCGA colon samples. Forty-two out of 53 CpGs were available for promoter region (left panel), while 62 out of 86 CpGs were available for gene body (right panel) in TCGA datasets. Salmon color represents normal colon tissues (n=38); cyan color represents colon tumors (n=292). Blue

indicates low beta values; red represents high beta values. (C) Supervised clustering of TCGA colon samples based on expression level of genes corresponding to differentially methylated CpGs. Separate heatmap showing expression level of 45 genes corresponding to the sites in promoter region (left panel), and 67 genes in the gene body region (right panel). Blue indicates low expression level; red represents high expression level.

Genes with increased methylation are enriched in cancer-related pathways and are NFAT, LEF1, MAZ-regulated

We next investigated the functions of the genes corresponding to the identified CpGs using the GeneCodis (v2) gene set enrichment analysis tool. Gene ontology enrichment analysis revealed that five out of the 10 (50%) most significant GO processes (FDR = 0.05) were related to transcription regulation, which is one of the key functional role of DNA methylation in order to control gene expression (Figure 5A). Notably, the list of enriched genes included well-known oncogenes, such as AFF3, CTNND2, ELK4, ESR1, PAX3, TRRAP, and WHSC1L1. The pathway enrichment analysis revealed that olfactory transduction and p53 signaling pathway were overrepresented (FDR= 0.05) in the gene set with increased methylation (Figure 5B). Enrichment analysis further revealed that the corresponding genes were enriched for alternating splicing as the major keyword (fold enrichment = 1.27, FDR = 0.000132, in Figure 5C).

Further, enrichment analysis among the transcription factors regulating these identified genes revealed that 47 (12%) of genes were regulated by the nuclear factor of activated T-cells (NFAT, $P = 1.48 \times 10^{-8}$, Supplementary Figure 3), 59 (15%) genes were regulated by lymphoid enhancer-binding factor 1 (LEF1) ($P = 2.1 \times 10^{-8}$, Figure 5D), and 51 (13%) genes by MYC Associated Zinc Finger Protein (MAZ) ($P = 4.07 \times 10^{-8}$, Figure 5E). Therefore, by increasing DNA methylation, decitabine strongly affected several well-known oncogenes related to cancer-related

pathways, especially the olfactory pathway and p53 tumor suppressor pathway, and a majority of these genes were regulated by NFAT, LEF1 and MAZ transcription factor (Table 1).

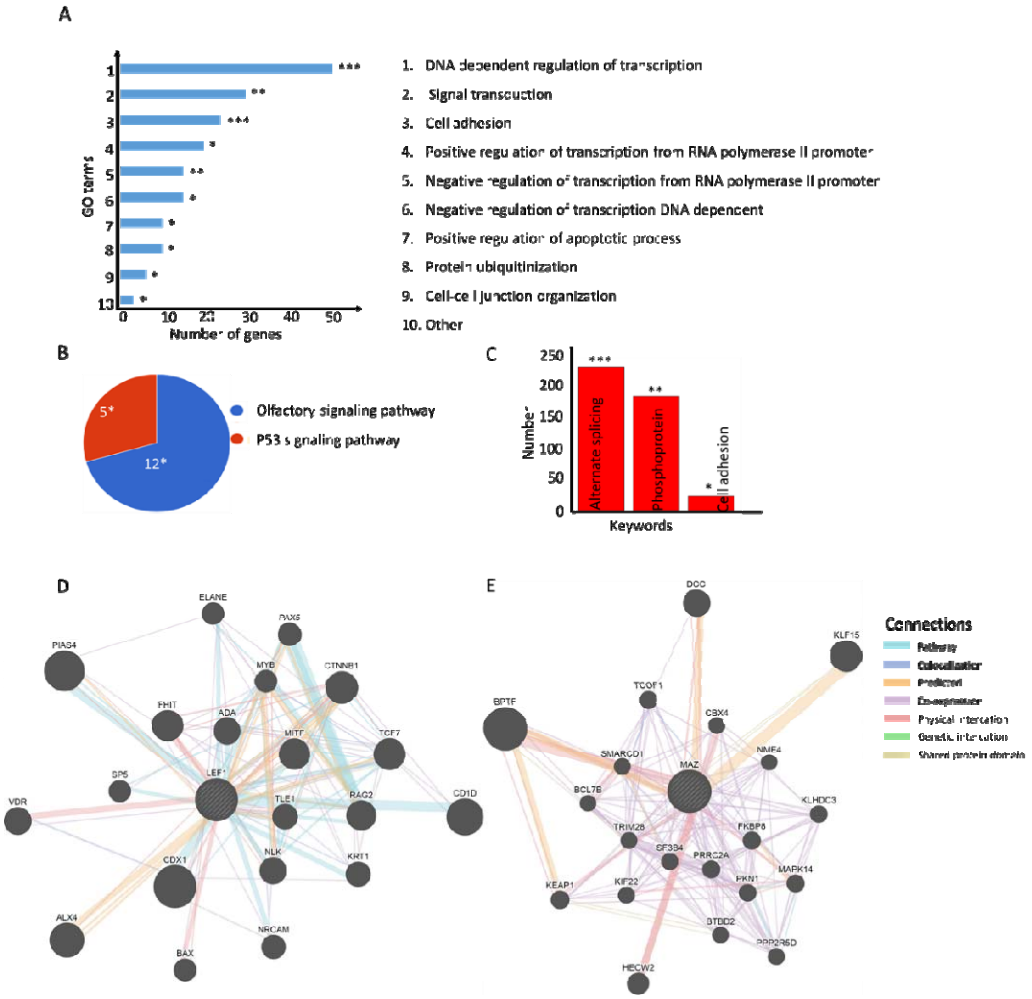


Figure 5: Gene ontology (GO) and pathway enrichment analyses of genes corresponding to differentially methylated CpGs.

(A) Top ten significantly enriched cellular processes have been shown as bar plot. The lengths of the bars denote the number of genes present in each of the top GO categories. (B) Pie-chart showing the significantly enriched pathways for the genes. The number of genes present in each pathway group has been shown along with the hypergeometric test p-value corrected for multiple testing. (C) The keywords enrichment analysis for the genes has been shown as bar chart. The length of the bar represents the number of genes enriched for each keyword, the FDR-corrected p-

*value has been shown at the top. (D) Interaction network of genes regulated by LEF1 (E) Interaction network of genes regulated by MAZ. Only interactions among those genes that are directly connected to enriched transcription factors (LEF1 and MAZ) have been shown as network using GENEMANIA. The size of the nodes is proportional to score calculated by GENEMANIA using label propagation algorithm that indicates the relevance of each gene to the original list of genes based on the selected networks. *P<0.05, **P<0.005, ***P<0.0005*

Discussion

Our study indicates that clinically feasible dose of decitabine (0.06 µM to 300 µM) treatment causes a transient increase in DNA methylation level of a small fraction of CpGs in the genome related to critical signaling pathways involved in tumorigenesis. The use of 3-day exposure with such doses *in vitro* produces a quick increase in DNA methylation that may reflect the immediate response of cells to external stimuli. However, increase in DNA methylation in most of the sites were not correlated with the corresponding gene expression levels, instead, they were enriched for alternative splicing as the key process term (Supplementary figure 4). Previous results suggest a complex nature of DNMTi action over cancer cells by not only changing the expression of certain genes, but also regulating the number of different transcript isoforms for several other genes [27]. One of the major role of DNA methylation is to regulate alternative splicing in the genome mainly by modulation of the elongation rate of RNA polymerase II (Pol II) by CCCTC-binding factor (CTCF) and methy-l-CpG binding protein 2 (MeCP2) [28]. Increase in DNA methylation can also enhance alternating splicing by the formation of a protein bridge by heterochromatin protein 1 (HP1) that recruits splicing factors onto transcribed alternative exons [29].

We believe that the transient increase in methylation due to decitabine treatment is mainly for the temporary alteration of transcript level of certain genes rather than permanent shut down or enhancement of expression as the increase in methylation is transient (vanishes 10 days after treatment). However, further study is needed to understand how the induced methylation due to DNMTi treatment affects alternating splicing in cancer cells.

Further, genes corresponding to identified CpGs with increased methylation were enriched among cancer-related pathways (olfactory and p53 pathway), and are regulated by Nfat, Lef1 and Maz transcription factor. Lef1 is a known target gene of Wnt/β-catenin pathway and is upregulated in colonic carcinogenesis where Wnt-3A/beta-catenin signaling induces transcription from the LEF-1 promoter [30]. Knockdown of LEF1 inhibits colon cancer progression *in vitro* and *in vivo* [31]. Further, enriched p53 pathway and olfactory receptor signaling pathways are a hallmark of multiple cancer type, and are involved in cell proliferation, migration, and apoptosis of cancerous cells [32, 33, 34]. Olfactory receptor and related signaling activation inhibit cell proliferation and apoptosis in colorectal cancer cells [33]. Enrichment of genes among the cancer-related pathways involved in cell proliferation, differentiation, and apoptosis suggests a non-random, systematic selection of genes for a transient increase in methylation in order to carry out cancer-related biological process in cells.

Based on our analysis, there are multiple reasons to suggest that at least one key mechanism underlying the anti-tumor responses to DNMTi treatment may involve an increase in DNA methylation level of specific genes in cancer cells. First, we showed an increase in the pattern of DNA methylation in more than one type of cancer cells. Second, as defined for an epigenetic change, these sustained changes persist for significant periods of time (at least more than 5 days) after a transient, subsequently withdrawn, drug exposure (in this case 72 hours). Third, the expression patterns for a subset of the genes are different between cancer and normal tissue types. Importantly, these changes are induced by drug doses that do not acutely kill cells and, thus, allow the transient alterations in gene methylation patterns to act on emerging molecular changes to cells after DNMTi therapy. We showed that these changes include anti-tumor events in multiple key pathways, such as p53 pathway, olfactory receptor pathway, regulation of transcription level, and others which can cause huge molecular cascade in the cells even after removal of the drug. Thus, increased methylation might be considered as a key feature of DNMTi therapy that can alter multiple cancer-related pathways simultaneously.

Conclusion

In summary, our findings provide novel insights into understanding the mechanism of action of DNMTi treatment in case of multiple cancer types, primarily colon cancer. Our findings suggest the existence of CpG sites in the genome that can resist DNMTi treatment and show an opposite effect of hypermethylation than expected demethylation, hence these CpGs could be clinically applicable as a response-predictive biomarker for patient stratification. Our results also suggest that DNMTi has a complex mechanism of action and a generalized pattern for the activity of DNMTi is challenging to find. Hence, the effects of DNMTi on cancer tissues should be analyzed at the individual gene level, rather than at the entire genomic level, and separately for each tissue type and even cancer patient.

Future perspective

Our analysis is the first of its kind that directly shows increased methylation level at certain loci after DNMTi treatment in the HCT116 genome, contrary to its classical well-studied role in decreasing methylation. These findings were also validated across multiple cells lines belonging to different cancer and tissue types. However, a more functional mechanistic study in higher model systems and human tissue types is required for revealing how the increase in methylation at individual loci alters treatment response and the pathological burden of disease. We hope that these current observations will have implications for further research about DNMTi response, as well as resistance mechanisms, with the aim use drug-induced methylation with DNMT inhibitors as tool for treatment strategy for multiple cancers.

Summary points

- DNMTi treatment increases DNA methylation in a small fraction of loci in HCT116 cells.
- The increase in methylation is transient and exists between 24 hours to at least 5 days.
- There is a limited correlation between DNA methylation of CpGs with increased methylation and gene expression and most of the genes with such CpGs are enriched for alternating splicing.
- A subset of the CpGs with increased methylation is differentially expressed between cancer and healthy tissue in the TCGA colon cancer data.

- 77% of genes having CpGs with increased methylations showed differential expression between the colon and normal tissues.
- Identified CpGs sites are enriched for enhancer regions in the genome.
- Genes corresponding to differentially methylated CpGs are enriched for p53 and olfactory receptor signaling pathways and are involved in transcriptional regulation of cells.
- These data suggest a complex nature of decitabine action on the genome and its effect need to be analyzed in a specific genetic context, instead of using pan-genome analysis.

References

1. Shenker N, Flanagan JM. Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *Br J Cancer* 106(2), 248-253 (2012).

2. Koch A, Joosten SC, Feng Z *et al*. Analysis of DNA methylation in cancer: location revisited. *Nat Rev Clin Oncol* 15(7), 459-466 (2018).

3. Ramos MP, Wijetunga NA, Mclellan AS, Suzuki M, Greally JM. DNA demethylation by 5-aza-2'-deoxycytidine is imprinted, targeted to euchromatin, and has limited transcriptional consequences. *Epigenetics Chromatin* 8 11 (2015).

4. Minkovsky A, Sahakyan A, Bonora G *et al*. A high-throughput screen of inactive X chromosome reactivation identifies the enhancement of DNA demethylation by 5-aza-2'-dC upon inhibition of ribonucleotide reductase. *Epigenetics Chromatin* 8 42 (2015).

5. Bohl SR, Bullinger L, Rucker FG. Epigenetic therapy: azacytidine and decitabine in acute myeloid leukemia. *Expert Rev Hematol* 11(5), 361-371 (2018).

6. Derissen EJ, Beijnen JH, Schellens JH. Concise drug review: azacitidine and decitabine. *Oncologist* 18(5), 619-624 (2013).

7. Sarkar S, Goldgar S, Byler S, Rosenthal S, Heerboth S. Demethylation and re-expression of epigenetically silenced tumor suppressor genes: sensitization of cancer cells by combination therapy. *EPIGENOMICS* 5(1), 87–94 (2013).

8. Kastl L, Brown I, Schofield AC. Altered DNA methylation is associated with docetaxel resistance in human breast cancer cells. *International Journal of Oncology* 36(5), (2010).

    - **The first paper that showed that decitabine treatment can increase DNMT level in docetaxel resistance MCF7 cells.**

9. Chowdhury B, Mcgovern A, Cui Y *et al*. The hypomethylating agent Decitabine causes a paradoxical increase in 5-hydroxymethylcytosine in human leukemia cells. *Sci Rep* 5 9281 (2015).

- **Shows evidence that decitabine treatment increases 5-hydroxymethylcytosine level, an oxidation product of methylcytosine.**

10. Broday L, Lee YW, Costa M. 5-azacytidine induces transgene silencing by DNA methylation in Chinese hamster cells. *Mol Cell Biol.19 (4)* 3198-204 (1999)

- **Shows evidence that decitabine analog, 5-azacytidine treatment also can increase methylation of a foreign gene in Chinese hamster cells.**

11. Weber G, Shendure J, Tanenbaum DM, Church GM, Meyerson M. Identification of foreign gene sequences by transcript filtering against the human genome. *Nat Genet* 30(2), 141-142 (2002).

12. Tobiasson M, Abdulkadir H, Lennartsson A, *et al*.Comprehensive mapping of the effects of azacitidine on DNA methylation, repressive/permissive histone marks and gene expression in primary cells from patients with MDS and MDS-related disease. *Oncotarget*. 8(17):28812-28825 (2017).

13. Huidobro C, Urdinguio RG, Rodriguez RM *et al*. A DNA methylation signature associated with aberrant promoter DNA hypermethylation of DNMT3B in human colorectal cancer. *Eur J Cancer* 48(14), 2270-2281 (2012).

14. De Carvalho DD, Sharma S, You JS *et al*. DNA methylation screening identifies driver epigenetic events of cancer cell survival. *Cancer Cell* 21(5), 655-667 (2012).

15. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* 26(4), 577-590 (2014).

16. Han H, Yang X, Pandiyan K, Liang G. Synergistic re-activation of epigenetically silenced genes by combinatorial inhibition of DNMTs and LSD1 in cancer cells. *PLoS One* 8(9), e75136 (2013).

17. Leadem BR, Kagiampakis I, Wilson C *et al*. A KDM5 Inhibitor Increases Global H3K4 Trimethylation Occupancy and Enhances the Biological Efficacy of 5-Aza-2'-Deoxycytidine. *Cancer Res* 78(5), 1127-1139 (2018).

18. Li H, Chiappinelli KB, Guzzetta AA, *et al*. Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. Oncotarget, 5(3):587-98, (2014).

- **Change in methylation of genes after azacytidine treatment have been investigated in 63 breast, ovarian and colon cancer cell lines.**

19. Chen YA, Lemire M, Choufani S *et al*. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 8(2), 203-209 (2013).

20. Teschendorff AE, Marabita F, Lechner M *et al*. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 29(2), 189-196 (2013).

21. Giri AK, Bharadwaj S, Banerjee P *et al*. DNA methylation profiling reveals the presence of population-specific signatures correlating with phenotypic characteristics. *Mol Genet Genomics* 292(3), 655-662 (2017).

22. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24(13), 1547-1548 (2008).

23. Ritchie ME, Phipson B, Wu D *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7), e47 (2015).

24. Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res* 40(Web Server issue), W478-483 (2012).

25. Zuberi K, Franz M, Rodriguez H *et al*. GeneMANIA prediction server 2013 update. *Nucleic Acids Res* 41(Web Server issue), W115-122 (2013).

26. Huang Da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1), 44-57 (2009).

27. Ding XL, Yang X, Liang G, Wang K. Isoform switching and exon skipping induced by the DNA methylation inhibitor 5-Aza-2'-deoxycytidine. *Sci Rep* 6 24545 (2016).

    - **Shows the effect of decitabine treatment in isoform switching and exon skipping during tranlation in UM-UC-3 (bladder cancer) cell line .**

28. Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. *Trends Genet* 31(5), 274-280 (2015).

    - **Discusses the role of DNA methylation in alternating splicing.**

29. Yearim A, Gelfman S, Shayevitch R *et al*. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep* 10(7), 1122-1134 (2015).

30. Kriegl L, Horst D, Reiche JA, Engel J, Kirchner T, Jung A. LEF-1 and TCF4 expression correlate inversely with survival in colorectal cancer. *J Transl Med* 8 123 (2010).

31. Wang WJ, Yao Y, Jiang LL *et al.* Knockdown of lymphoid enhancer factor 1 inhibits colon cancer progression in vitro and in vivo. *PLoS One* 8(10), e76596 (2013).

32. Li XL, Zhou J, Chen ZR, Chng WJ. P53 mutations in colorectal cancer - molecular pathogenesis and pharmacological reactivation. *World J Gastroenterol* 21(1), 84-93 (2015).

33. Weber L, Al-Refae K, Ebbert J *et al*. Activation of odorant receptor in colorectal cancer cells leads to inhibition of cell proliferation and apoptosis. *PLoS One* 12(3), e0172491 (2017).

- **Shows the role of olfactory receptor pathway in colon cancer.**

34. Weber L, Massberg D, Becker C *et al*. Olfactory Receptors as Biomarkers in Human Breast Carcinoma Tissues. *Front Oncol* 8 33 (2018).

Table 1: Top 10 transcription factor enriched in gene set corresponding to induced CpGs and its regulated genes

S.N.	Transcription factor	Number of genes	FDR	Regulated genes
1	NFAT	47	1.48×10^{-8}	SLIT3,RORA, SYT10,CNTNAP2,SCN3A, MRPL28,ADAMTSL1,CTNND2,ESR1,PPM1B, PDE4D,RARB, OPCML,SNX15,PRKD2,ID3,PTPRO, ADCY2,CUL3,DMD,ITM2C,KLF12,BCOR, CTNND1,SGCD,ACACA,HDAC6,POGK,AUTS2,PAX3,DLG2,SLC6A5,SOX5,DLC1,ANTXR1,NGFRAP1,LSAMP,GRM8,CACNA2D3,ETS1,S100A10,ADAMTS17,KCNH5,ARHGAP6,KCNMA1,MAP7,KCNN2
2	LEF1	59	2.10×10^{-8}	PDCD10,NRXN1,SLIT3,RORA, YWHAZ, COX7B, SYNPR, SCN3A, SORCS1,TMSB4X,ADAMTSL1,NXN, CLSTN2,ZNF8,CNTN6,MAGED2,WHSC1L1,GMPR2,PDE4D,ABCF2,RARB,CNKSR2,TIA1,SMARCA1,SFRP2,OPCML, WDFY3,MBTD1,CACNA1E,PTPRO,MCTS1,DMD,KLF12,CTNND1,SGCD,ACACA,POGK,OXCT1,TAF1,PAX3,SLC6A5,SOX5,CPS1,SIX4,GPC6,DLC1,GTF2A2,TLE3,GAB2,TMSL3,CACNA2D3,ETS1,BZW1,ADAMTS12,CD160,TCERG1L, KCNH5,ARHGAP6,NR2F1
3	MAZ	51	4.07×10^{-8}	PDCD10,SLIT3,SV2B, RORA, YWHAZ, CNTNAP2,SORCS1,PRKCI,CTNND2,ACCN2,ESR1,MAGED2,HRK, FKBP2,RARB, CNKSR2,SSR1,SMARCA1,SFRP2,PRKD2,UBE2L3,PTPRF, ID3,DMD,ITM2C,KLF12,P4HA1,BCOR, RGS7,CTNND1,DUSP6,POGK, TAF1,LYPLA2,AUTS2,DLG2,SOX5,SIX4,DLC1,TLE3,PRDM16,NGFRAP1,POLR1D,ARVCF, BZW1,THRAP3,TRRAP, ZNRF1,KCNH5,KCNMA1,NR2F1
4	OCT1	16	1.49×10^{-7}	NRXN1,NR2C2,SCN3A, PDE4D,RARB, DMD,KLF12,DUSP6,SOX5,DLC1,TLE3,GAB2,TMSL3,TCERG1L,IRX2,NR2F1
5	MEF2	27	2.32×10^{-7}	HIPK1,CTNNA1,NRXN3,ESR1,PPM1B, SLC12A1,SMARCA1,OPCML,WBP5,ADCY2,CUL3,DMD,KLF12,P4HA1,CTNND1,SGCD,DLG2,SOX5,GLG1,NELL2,GRM8,CACNA2D3,ETS1,CIAO1,ADAMTS12,AGTPBP1,NR2F1
6	CHX10	27	4.22×10^{-7}	RORA, CTNND2,CLSTN2,ACCN2,HRK, RARB, PTPRO, MCTS1,DMD,CTNND1,DIXDC1,PAX3,DLG2,SLC6A5,SOX5,DPP10,DLC1,NELL2,TMSL3,GRM8,CACNA2D3,ADAMTS17,KCNH5,ARHGAP6,IRX2,KCNMA1,NR2F1
7	GFI1	20	7.37×10^{-7}	NRXN1,ICAM5,RORA, NRXN3,CLSTN2,ESR1,CNTN6,PDE4D,ABCF2,RARB, CACNA1E, MCTS1,DMD,KLF12,CTNND1,SLC6A5,SOX5,CRYZL1,LSAMP, ETS1
8	OCT1	15	8.19×10^{-7}	NRXN1,SYNPR, NRXN3,PDE4D,SFRP2,ID3,DMD,KLF12,BCOR, DUSP6,SOX5,GAB2,CRYZL1,TMSL3,IRX2
9	OCT	19	1.95×10^{-6}	NRXN1,SYNPR, NRXN3,PDE4D,RARB, SMARCA1,SFRP2,WBP5,CUL3,DMD,BCOR, SOX5,TLE3,CNOT2,GAB2,NGFRAP1,CRYZL1,TMSL3,TCERG1L
10	OCT1	14	5.60×10^{-6}	NRXN1,NR2C2,SYNPR, NRXN3,PDE4D,SFRP2,ID3,DMD,KLF12,DUSP6,DLG2,GAB2,CRYZL1,TMSL3

FDR-corrected p-values from the hypergeometric test are shown in the table.

Supplementary Table 1: List of 638 CpG showing increased methylation after decitabine treatment in HCTK116 cell line

SN	Chromosome	Base position	CpG	Position in gene	Gene
1	5	76332945	ch.5.1443044F	Body	AGGF1
2	10	79793495	cg23620279	Promoter	RPS24
3	12	58087540	cg18202167	Promoter	OS9
4	12	58087544	cg00062356	Promoter	OS9
5	7	72299837	cg19955956	Promoter	SBDSP
6	2	231729487	cg15127563	Promoter	ITM2C
7	15	70390363	cg07020846	Promoter	TLE3
8	6	33282175	cg06097707	Promoter	TAPBP
9	8	17942583	cg07312099	Promoter	ASAH1
10	X	53310991	cg14972383	Promoter	IQSEC2
11	7	150924351	cg10436877	Promoter	ABCF2
12	X	12993075	cg23376554	Promoter	TMSB4X
13	X	102510216	cg14004049	Promoter	TCEAL8
14	1	43996436	cg14172596	Promoter	PTPRF
15	1	113933413	cg18539474	Promoter	MAGI3
16	7	107221074	cg11785538	Promoter	BCAP29
17	13	49684397	cg17091793	Promoter	FNDC3A
18	5	78281964	cg26802063	Promoter	ARSB
19	X	146312384	cg27167381	Promoter	MIR506
20	X	77154874	cg10646076	Promoter	COX7B
21	4	7105115	cg26600181	Promoter	FLJ36777
22	10	91152122	cg16395953	Promoter	IFIT1
23	X	101186742	cg23922730	Promoter	ZMAT1
24	4	47033180	cg21472546	Promoter	GABRB1
25	1	1981816	cg22865720	Promoter	PRKCZ
26	11	55587104	cg08060810	Promoter	OR5D18
27	10	63661280	cg16253809	Promoter	ARID5B
28	1	23886472	cg20485144	Promoter	ID3
29	19	58790298	cg23548487	Promoter	ZNF8

30	X	145082826	cg15918587	Promoter	MIR891B
31	19	13885098	cg09952620	Promoter	C19orf53
32	11	5905350	cg06484232	Promoter	OR52E4
33	13	103426305	cg13675958	Promoter	C13orf27
34	16	5008134	cg08043782	Promoter	SEC14L5
35	5	158634905	cg26362852	Promoter	RNF145
36	19	20368371	cg22155405	Promoter	LOC284441
37	19	4066818	cg10561472	Promoter	ZBTB7A
38	6	13615538	cg05884522	Promoter	NOL7
39	2	39005241	cg15934678	Promoter	GEMIN6
40	5	92918848	cg05945291	Promoter	NR2F1
41	X	2984799	cg17012513	Promoter	ARSF
42	10	15210836	cg23193446	Promoter	NMT2
43	19	20011538	cg27379065	Promoter	ZNF93
44	13	28194831	cg07375367	Promoter	POLR1D
45	1	225616668	cg20483690	Promoter	LBR
46	X	119737675	cg05782751	Promoter	MCTS1
47	19	54694174	cg12173535	Promoter	MBOAT7
48	19	37020332	cg24909706	Promoter	ZNF260
49	11	131779469	cg08097520	Promoter	NTM
50	15	72524656	cg25016070	Promoter	PKM2
51	4	25162716	cg19113954	Promoter	SEPSECS
52	5	92918072	cg16448525	Promoter	FLJ42709
53	6	160211006	cg07151830	Promoter	TCP1
54	5	135702333	cg16219583	Promoter	TRPC7
55	5	135701422	cg17275074	Promoter	TRPC7
56	8	24240597	cg14143055	Promoter	ADAMDEC1
57	1	3606550	cg21388339	Promoter	TP73
58	X	153169465	cg20664654	Promoter	AVPR2
59	8	28748404	cg06241765	Promoter	INTS9

60	X	102629870	cg13486082	Promoter	NGFRAP1
61	5	92917334	cg08003613	Promoter	FLJ42709
62	2	70476188	cg24074685	Promoter	TIA1
63	X	102629912	cg25198830	Promoter	NGFRAP1
64	X	12992684	cg17625764	Promoter	TMSL3
65	19	30432806	cg27180365	Promoter	C19orf2
66	12	50450889	cg23126949	Promoter	ACCN2
67	13	34391699	cg08155354	Promoter	RFC3
68	14	75535757	cg15979150	Promoter	FAM164C
69	3	69129729	cg21635584	Promoter	UBA3
70	1	25559743	cg17651255	Promoter	SYF2
71	9	113800909	cg06148685	Promoter	LPAR1
72	12	19282261	cg13108328	Promoter	PLEKHA5
73	18	48404491	cg26727372	Promoter	ME2
74	X	128657727	cg18959966	Promoter	SMARCA1
75	1	10532838	cg21149582	Promoter	DFFA
76	X	15354150	cg18016370	Promoter	PIGA
77	4	85888001	cg14553853	Promoter	WDFY3
78	11	108092818	cg12019961	Promoter	ATM
79	20	20032580	cg06599170	Promoter	C20orf26
80	4	68567439	cg25695041	Promoter	UBA6
81	3	52719268	cg09817993	Promoter	GNL3
82	6	30655567	cg23903723	Promoter	KIAA1949
83	9	139376822	cg13820039	Promoter	C9orf163
84	11	79114133	cg25837979	Promoter	MIR708
85	5	149379518	cg26588194	Promoter	HMGXB3
86	1	145469887	cg12222699	Promoter	ANKRD34A
87	6	3458177	cg11713788	Promoter	SLC22A23
88	19	12625436	cg24109012	Promoter	ZNF709
89	13	29597447	cg07790085	Promoter	MTUS2

90	11	5019849	cg23434090	Promoter	OR51L1
91	14	61191253	cg09970023	Promoter	SIX4
92	1	234040045	cg10878114	Promoter	SLC35F3
93	5	54319373	cg14597388	Promoter	GZMK
94	11	51413644	cg23935054	Promoter	OR4A5
95	16	420755	cg09504571	Promoter	MRPL28
96	19	20749738	cg12124647	Promoter	ZNF737
97	10	13628544	cg15881990	Promoter	PRPF18
98	10	15902872	cg19707359	Promoter	FAM188A
99	1	152087267	cg22603037	Promoter	TCHH
100	5	140592867	cg14640659	Promoter	PCDHB13
101	11	111807548	cg11471799	Promoter	DIXDC1
102	11	128458153	cg16792062	Promoter	ETS1
103	11	55577775	cg05788138	Promoter	OR5L1
104	14	78869352	cg10828316	Promoter	NRXN3
105	12	113773298	cg12583184	Promoter	SLC24A6
106	X	77151316	cg11290168	Promoter	MAGT1
107	15	102344469	cg26155802	Promoter	OR4F6
108	X	154254823	cg09526164	Promoter	FUNDC2
109	1	16010601	cg09073052	Promoter	PLEKHM2
110	11	55796572	cg17060964	Promoter	OR5AS1
111	15	54303927	cg22845496	Promoter	UNC13C
112	11	123814972	cg15625631	Promoter	OR6T1
113	4	184365198	cg24787081	Promoter	CDKN2AIP
114	1	159258877	cg14696870	Promoter	FCER1A
115	19	14683110	cg16256643	Promoter	NDUFB7
116	22	20004611	cg12912949	Promoter	ARVCF
117	1	151967449	cg06698332	Promoter	S100A10
118	X	99892000	cg11509733	Promoter	TSPAN6
119	14	70715730	cg20576094	Promoter	ADAM21P1

120	1	241694605	cg11150901	Promoter	KMO
121	5	68513219	cg15387943	Promoter	MRPS36
122	5	159846543	cg15333689	Promoter	SLU7
123	14	55595666	cg26335127	Promoter	LGALS3
124	X	77154732	cg24112882	Promoter	COX7B
125	4	155471778	cg11404039	Promoter	PLRG1
126	6	52149972	cg18225895	Promoter	MCM3
127	12	117319577	cg14276619	Promoter	HRK
128	10	44287023	cg17952824	Promoter	HNRNPA3P1
129	12	67662516	cg18750937	Promoter	CAND1
130	1	248568331	cg13053563	Promoter	OR2T1
131	17	3627058	cg14971744	Body	ITGAE
132	15	89010209	cg20630605	Body	MRPL46
133	12	97885270	cg27533635	Body	RMST
134	X	131547702	cg13633856	Body	MBNL3
135	X	131547697	cg14520512	Body	MBNL3
136	8	12974556	cg06103928	Body	DLC1
137	4	76439140	cg12093136	Body	RCHY1
138	5	54455564	cg08212230	Body	CDC20B
139	21	39494547	cg25816610	Body	DSCR8
140	1	231820076	cg07134368	Body	TSNAX-DISC1
141	1	231964048	cg22367981	Body	DISC1
142	1	44287964	cg23290313	Body	ST3GAL3
143	3	171138553	cg22901347	Body	TNIK
144	2	223151884	cg11490745	Body	PAX3
145	16	74565916	ch.16.1684049R	Body	GLG1
146	12	45034784	ch.12.897509F	Body	NELL2
147	19	52862167	cg10341573	Body	ZNF610
148	1	114503218	ch.1.2681285F	Body	HIPK1
149	10	79033545	cg08772567	Body	KCNMA1

150	14	33826344	cg15454195	Body	NPAS3
151	19	53619086	cg22834281	Body	ZNF415
152	X	21559778	ch.X.346519R	Body	CNKSR2
153	20	45937282	ch.20.1002962F	Body	ZMYND8
154	7	69447465	cg09703727	Body	AUTS2
155	5	155909246	cg24132325	Body	SGCD
156	2	205591269	cg18626478	Body	PARD3B
157	X	70661061	ch.X.1084407R	Body	TAF1
158	10	83848597	cg17519477	Body	NRG3
159	3	189353419	cg25708695	Body	TP63
160	22	21968010	ch.22.149158R	Body	UBE2L3
161	14	80324276	cg19753609	Body	NRXN3
162	7	69478390	cg16819888	Body	AUTS2
163	7	154006066	cg07467482	Body	DPP6
164	2	165998136	cg16631432	Body	SCN3A
165	5	156097036	cg15160274	Body	SGCD
166	2	50570407	cg06707406	Body	NRXN1
167	20	34297200	ch.20.707667F	Body	RBM39
168	7	147709862	cg22807241	Body	MIR548F3
169	11	115369647	cg11019127	Body	CADM1
170	3	25635650	cg07405178	Body	RARB
171	10	108674143	cg23024358	Body	SORCS1
172	4	85766242	ch.4.1647744F	Body	WDFY3
173	2	100371023	cg22092126	Body	AFF3
174	2	153575717	cg07491444	Body	ARL6IP6
175	13	34392781	cg24492140	Body	RFC3
176	2	159173778	cg21514997	Body	CCDC148
177	12	23998997	cg06764736	Body	SOX5
178	13	43545157	cg17400905	Body	EPSTI1
179	11	115096810	cg25461513	Body	CADM1

180	19	19923956	ch.19.841535R	Body	ZNF506
181	1	28886386	ch.1.953398R	Body	TRNAU1AP
182	2	116480197	cg06394103	Body	DPP10
183	5	59126518	cg27583655	Body	PDE4D
184	22	46114168	ch.22.909671F	Body	ATXN10
185	1	16715418	ch.1.572291F	Body	C1orf144
186	5	113805552	cg25486361	Body	KCNN2
187	13	60543691	ch.13.865492R	Body	DIAPH3
188	2	100175805	cg17165836	Body	AFF3
189	2	100365075	cg13361307	Body	AFF3
190	15	76136846	cg16242106	Body	UBE2Q2
191	12	50635579	ch.12.1023240F	Body	LIMA1
192	2	227850069	cg09157320	Body	RHBDD1
193	20	10026325	ch.20.221631R	Body	ANKRD5
194	11	64795449	cg21821990	Body	SNX15
195	12	117611422	ch.12.2406115F	Body	FBXO21
196	14	63508497	cg25609301	Body	KCNH5
197	8	38174205	ch.8.903080R	Body	WHSC1L1
198	2	192543258	cg15794798	Body	OBFC2A
199	7	150452495	cg14319487	Body	LOC100128542
200	15	61346347	cg08099431	Body	RORA
201	15	24411878	cg15564871	Body	PWRN2
202	10	132942686	cg06938601	Body	TCERG1L
203	13	26483520	cg12565580	Body	ATP8A2
204	3	194136354	ch.3.3822654R	Body	ATP13A3
205	5	11529629	cg16051561	Body	CTNND2
206	6	1836850	cg21478123	Body	GMDS
207	16	75044269	ch.16.1700675R	Body	ZNRF1
208	10	132942731	cg25486749	Body	TCERG1L
209	5	166938213	cg23167425	Body	ODZ2

210	11	92264986	cg07276831	Body	FAT3
211	18	34194679	ch.18.672159R	Body	FHOD3
212	16	9010914	ch.16.350833F	Body	USP7
213	3	130299763	cg17937340	Body	COL6A6
214	6	1909853	cg11276500	Body	GMDS
215	10	15901893	cg19053479	Body	FAM188A
216	7	48319696	cg10626169	Body	ABCA13
217	1	45187551	cg07722722	Body	C1orf228
218	3	140229290	cg23463099	Body	CLSTN2
219	10	96306185	cg10069677	Body	HELLS
220	11	122037845	cg15826891	Body	LOC399959
221	5	168395173	cg11458498	Body	SLIT3
222	21	34978286	cg12708807	Body	CRYZL1
223	5	171653553	ch.5.3268483F	Body	UBTD2
224	15	80188894	cg07293993	Body	MTHFS
225	12	50527085	ch.12.1019410F	Body	LASS5
226	1	10464086	ch.1.385573R	Body	PGD
227	3	21558209	cg16439360	Body	ZNF385D
228	1	241474238	cg08238568	Body	RGS7
229	5	138146102	ch.5.2559743R	Body	CTNNA1
230	5	41837108	ch.5.884579R	Body	OXCT1
231	10	114074843	cg18014500	Body	GUCY2G
232	1	181514216	cg22359828	Body	CACNA1E
233	17	1717862	ch.17.79071R	Body	SMYD4
234	5	7686199	cg16253976	Body	ADCY2
235	2	225441832	cg11229715	Body	CUL3
236	14	91751397	ch.14.1452150F	Body	CCDC88C
237	8	35401908	cg22872195	Body	UNC5D
238	8	97855859	cg08247527	Body	PGCP
239	10	34817409	cg19017553	Body	PARD3

240	10	106749483	cg26349484	Body	SORCS3
241	3	35730993	cg15459537	Body	ARPP-21
242	9	14838752	cg13762569	Body	FREM1
243	15	41734226	ch.15.433532F	Body	RTF1
244	6	7297596	ch.6.197209F	Body	SSR1
245	13	94493055	cg21222888	Body	GPC6
246	8	119282796	ch.8.2353618R	Body	SAMD12
247	5	15780558	ch.5.409282R	Body	FBXL7
248	4	72635202	cg24806812	Body	GC
249	2	211401919	ch.2.4215183F	Body	CPS1
250	3	64801329	cg21324884	Body	MIR548A2
251	4	110553306	ch.4.2065340F	Body	CCDC109B
252	2	105686129	ch.2.2207852R	Body	MRPS9
253	1	212238115	ch.1.4129519F	Body	DTL
254	6	12910610	cg14773588	Body	PHACTR1
255	5	118874673	ch.5.2173511R	Body	HSD17B4
256	7	150451091	cg06276978	Body	LOC100128542
257	20	29847402	cg25361651	Body	DEFB115
258	3	54807076	cg19093405	Body	CACNA2D3
259	1	206768238	ch.1.4018176R	Body	LGTN
260	1	247691111	cg25139877	Body	LOC148824
261	4	44425078	cg12709692	Body	KCTD8
262	3	45553017	cg15691035	Body	LARS2
263	13	74491964	ch.13.1085822R	Body	KLF12
264	7	148111040	ch.7.3089487R	Body	CNTNAP2
265	3	121977827	cg10364968	Body	CASR
266	6	167200499	cg18495191	Body	RPS6KA2
267	17	49281558	ch.17.1348593F	Body	MBTD1
268	1	21074008	ch.1.705736F	Body	HP1BP3
269	18	55024674	cg21518865	Body	ST8SIA3

270	5	33698653	ch.5.731560F	Body	ADAMTS12
271	10	69913749	cg18986048	Body	MYPN
272	2	1182847	ch.2.35699F	Body	SNTG2
273	8	144798631	cg22892110	Body	MAPK15
274	15	48515109	cg23530596	Body	SLC12A1
275	6	33393183	cg11261678	Body	SYNGAP1
276	1	36766063	ch.1.1168472R	Body	THRAP3
277	6	33395430	cg19968421	Body	SYNGAP1
278	5	136682394	cg19567866	Body	SPOCK1
279	6	136828807	ch.6.2623783F	Body	MAP7
280	9	18825658	cg13724111	Body	ADAMTSL1
281	3	124726246	ch.3.2442921F	Body	HEG1
282	4	154707153	cg11467638	Body	SFRP2
283	3	51637060	ch.3.1119246R	Body	RAD54L2
284	5	160053412	cg27141889	Body	ATP10B
285	8	14108012	ch.8.362960F	Body	SGCZ
286	3	73591147	cg09236445	Body	PDZRN3
287	X	154255950	cg10432310	Body	FUNDC2
288	20	51648147	cg09566894	Body	TSHZ2
289	4	159918056	cg22266824	Body	C4orf45
290	2	44182106	ch.2.1056241F	Body	LRPPRC
291	9	111807413	ch.9.1678974F	Body	C9orf5
292	20	24953309	ch.20.532344R	Body	C20orf3
293	17	48194635	cg11441693	Body	SAMD14
294	14	79558823	cg18818949	Body	NRXN3
295	12	122065180	cg24082347	Body	ORAI1
296	5	98107521	cg12949466	Body	RGMB
297	1	241176676	cg22231602	Body	RGS7
298	19	12277357	cg13689563	Body	ZNF136
299	20	62645068	ch.20.1534602F	Body	PRPF6

300	15	61500965	cg20124735	Body	RORA
301	15	59940044	ch.15.825727F	Body	GTF2A2
302	8	13092547	ch.8.343778F	Body	DLC1
303	5	126871056	ch.5.2320326F	Body	PRRC1
304	11	47659584	ch.11.997072R	Body	MTCH2
305	11	20621341	cg20632573	Body	SLC6A5
306	17	837017	cg07494499	Body	NXN
307	4	121668750	ch.4.2245532F	Body	PRDM5
308	19	10401361	cg07097925	Body	ICAM5
309	3	169994002	ch.3.3303606F	Body	PRKCI
310	1	242310145	cg07977614	Body	PLD5
311	5	167028535	cg15651267	Body	ODZ2
312	3	15058168	ch.3.343413R	Body	NR2C2
313	8	36763165	cg12033248	Body	KCNU1
314	11	132931732	cg18413062	Body	OPCML
315	7	32038956	cg13298997	Body	PDE1C
316	9	88291443	ch.9.1152820R	Body	AGTPBP1
317	X	10087726	cg25497053	Body	WWC3
318	12	126055778	cg19992906	Body	TMEM132B
319	7	98536084	ch.7.2068158F	Body	TRRAP
320	4	128654094	cg13665890	Body	SLC25A31
321	7	79765394	cg25702790	Body	GNAI1
322	11	78071309	ch.11.1702122F	Body	GAB2
323	12	33589594	cg06721860	Body	SYT10
324	11	63140804	cg10116443	Body	SLC22A9
325	8	126142264	cg12803053	Body	NSMCE2
326	3	159590447	cg09811510	Body	SCHIP1
327	3	1382597	cg16522250	Body	CNTN6
328	3	115618688	cg16752940	Body	LSAMP
329	4	6676521	cg16758887	Body	LOC93622

330	11	78400028	cg07441953	Body	ODZ4
331	12	42877995	cg20908919	Other	PRICKLE1
332	4	130017238	cg13107060	Other	C4orf33
333	19	13262082	cg15340644	Other	IER2
334	5	59782121	cg18611813	Other	PDE4D
335	21	37433149	cg17039262	Other	SETD4
336	12	88536565	cg20328917	Other	TMTC3
337	12	121790694	cg12210527	Other	ANAPC5
338	6	36409495	cg06422757	Other	PXT1
339	1	34328907	cg05723953	Other	HMGB4
340	X	31889692	cg20522855	Other	DMD
341	11	83393062	cg13572369	Other	DLG2
342	19	37061383	cg10172250	Other	ZNF529
343	X	11282604	cg08456555	Other	ARHGAP6
344	6	153303350	cg18198306	Other	FBXO5
345	6	31515398	cg00002930	Other	NFKBIL1
346	X	54834954	cg09208571	Other	MAGED2
347	20	47778018	ch.20.1062061F	Other	STAU1
348	8	101963358	cg11839355	Other	YWHAZ
349	10	43902500	cg12891252	Other	HNRNPF
350	10	74855378	cg19832312	Other	P4HA1
351	X	39954231	cg07099245	Other	BCOR
352	18	3263082	cg19146448	Other	MYL12B
353	19	47219957	cg25522119	Other	PRKD2
354	9	74979581	cg14026485	Other	ZFAND5
355	2	208489667	cg07496861	Other	FAM119A
356	18	13375474	cg26700919	Other	C18orf1
357	3	132772692	cg24837219	Other	TMEM108
358	2	159905508	ch.2.3260358R	Other	TANC1
359	4	129731973	cg11264547	Other	PHF17

360	18	13375540	cg21243597	Other	C18orf1
361	3	114599007	cg19149693	Other	ZBTB20
362	15	91646263	cg27121538	Other	SV2B
363	X	50028082	cg11113650	Other	CCNB3
364	1	205599988	cg15558299	Other	ELK4
365	18	13229005	ch.18.316502R	Other	C18orf1
366	11	57529465	cg19210276	Other	CTNND1
367	22	29137759	cg22585269	Other	CHEK2
368	X	150867017	cg15731296	Other	PRRG3
369	19	38827331	cg09094448	Other	CATSPERG
370	6	152085641	cg18132851	Other	ESR1
371	X	101186679	cg26142661	Other	ZMAT1
372	12	24577904	cg24011341	Other	SOX5
373	1	26565342	ch.1.876374R	Other	CCDC21
374	5	147101975	cg11799006	Other	JAKMIP2
375	10	21807252	cg25195795	Other	C10orf140
376	2	16839610	cg09324018	Other	FAM49A
377	3	125093863	cg15705999	Other	ZNF148
378	3	176914208	cg07883762	Other	TBL1XR1
379	16	62067937	cg07244927	Other	CDH8
380	6	166581272	cg00070318	Other	T
381	1	158149974	cg24432768	Other	CD1D
382	4	87857667	cg19533294	Other	AFF1
383	12	87106229	cg14783993	Other	MGAT4C
384	11	85359560	cg26796873	Other	TMEM126A
385	1	33722623	cg11416597	Other	ZNF362
386	18	21976748	cg20528338	Other	OSBPL1A
387	2	64880293	cg11920737	Other	SERTAD2
388	X	48660813	cg10783042	Other	HDAC6
389	1	24118400	cg11659749	Other	LYPLA2

390	2	201677426	cg14211387	Other	BZW1
391	7	32526065	cg26856631	Other	LSM5
392	19	57324295	cg08155759	Other	PEG3
393	19	37001658	cg25505109	Other	ZNF260
394	15	100253379	ch.15.1787851R	Other	MEF2A
395	5	2746667	cg07766803	Other	IRX2
396	19	21608124	cg19023258	Other	ZNF493
397	3	63602009	cg11876912	Other	SYNPR
398	19	12299253	cg21880712	Other	ZNF136
399	15	20737822	cg14783259	Other	GOLGA6L6
400	19	53642858	cg15050103	Other	ZNF347
401	18	8638712	ch.18.189560F	Other	RAB12
402	19	20231820	cg06736434	Other	ZNF90
403	2	44460827	cg19769080	Other	PPM1B
404	10	78635553	cg23533270	Other	KCNMA1
405	7	117835958	cg13799581	Other	NAA38
406	6	25788879	cg06885175	Other	SLC17A1
407	10	32300362	ch.10.820670F	Other	KIF5B
408	X	7270088	cg10073470	Other	STS
409	4	6643382	cg20272423	Other	MRFAP1
410	2	96939558	ch.2.2007613R	Other	CIAO1
411	15	30930499	cg00067141	Other	ARHGAP11B
412	14	24701654	cg07519822	Other	GMPR2
413	15	99789637	cg25385940	Other	TTC23
414	X	80457315	cg21896142	Other	HMGN5
415	18	616707	cg23661343	Other	CLUL1
416	X	102611415	cg27464574	Other	WBP5
417	X	102611412	cg13208102	Other	WBP5
418	10	81107244	cg16098780	Other	PPIF
419	1	85528044	cg22488158	Other	WDR63

420	X	99899378	cg24666876	Other	SRPX2
421	X	77154996	cg15830530	Other	COX7B
422	20	43835661	cg11264863	Other	SEMG1
423	4	25235765	cg12931625	Other	PI4K2B
424	8	65496126	cg07205627	Other	BHLHE22
425	12	62654245	cg18915437	Other	USP15
426	8	145735102	cg17958180	Other	MFSD3
427	11	55606710	cg12962308	Other	OR5D16
428	6	32634362	cg05724777	Other	HLA-DQB1
429	3	149057820	cg06305422	Other	Intergeneic
430	3	152871396	ch.3.3016567F	Other	Intergeneic
431	2	19616327	cg21039221	Other	Intergeneic
432	1	209365350	cg22736624	Other	Intergeneic
433	3	116936427	cg11607648	Other	Intergeneic
434	8	89664399	ch.8.89733515F	Other	Intergeneic
435	22	32046895	ch.22.436090R	Other	Intergeneic
436	11	105098067	ch.11.104603277F	Other	Intergeneic
437	5	123635353	ch.5.2251785F	Other	Intergeneic
438	13	66774231	ch.13.65672232R	Other	Intergeneic
439	1	11967826	cg22340067	Other	Intergeneic
440	15	53205349	ch.15.50992641R	Other	Intergeneic
441	2	129663717	cg19404692	Other	Intergeneic
442	10	44099144	cg26270975	Other	Intergeneic
443	4	162116791	ch.4.162336241R	Other	Intergeneic
444	16	49903138	ch.16.48460639F	Other	Intergeneic
445	2	81088536	ch.2.80942047R	Other	Intergeneic
446	6	26595126	cg19497998	Other	Intergeneic
447	2	8397810	cg19256423	Other	Intergeneic
448	14	86554255	cg13090238	Other	Intergeneic
449	8	37006314	ch.8.870369R	Other	Intergeneic

450	6	128891095	cg18500322	Other	Intergeneic
451	1	161339556	ch.1.159606180R	Other	Intergeneic
452	12	85844287	ch.12.1700408R	Other	Intergeneic
453	8	54188341	ch.8.54350894R	Other	Intergeneic
454	14	53850070	ch.14.628538R	Other	Intergeneic
455	9	76690975	ch.9.919537F	Other	Intergeneic
456	2	227865079	ch.2.4543734F	Other	Intergeneic
457	11	81757203	ch.11.1767550R	Other	Intergeneic
458	18	5233979	cg26881207	Other	Intergeneic
459	X	33740873	cg07983986	Other	Intergeneic
460	2	205130274	cg17591195	Other	Intergeneic
461	8	119971687	cg21022303	Other	Intergeneic
462	2	118979739	cg27358426	Other	Intergeneic
463	11	15363270	cg09663343	Other	Intergeneic
464	8	2483325	cg17224775	Other	Intergeneic
465	5	178483871	cg11282433	Other	Intergeneic
466	2	157192128	cg12335829	Other	Intergeneic
467	3	137492929	cg15062059	Other	Intergeneic
468	3	16768672	ch.3.382096F	Other	Intergeneic
469	10	132834807	cg11315633	Other	Intergeneic
470	2	9188970	ch.2.246819F	Other	Intergeneic
471	15	75206153	ch.15.72993206F	Other	Intergeneic
472	3	71636085	cg06479142	Other	Intergeneic
473	14	54572859	cg24011936	Other	Intergeneic
474	6	164526833	cg21567971	Other	Intergeneic
475	2	181200445	ch.2.180908690F	Other	Intergeneic
476	16	13461010	ch.16.486323F	Other	Intergeneic
477	8	26942425	cg23760300	Other	Intergeneic
478	6	163768411	cg20867674	Other	Intergeneic
479	5	60527978	ch.5.1161320F	Other	Intergeneic

480	13	54817511	cg09602751	Other	Intergeneic
481	9	97015266	ch.9.96055087R	Other	Intergeneic
482	2	5374307	ch.2.154144R	Other	Intergeneic
483	20	11513716	ch.20.250771F	Other	Intergeneic
484	5	169008308	cg09106932	Other	Intergeneic
485	10	79401752	cg27024057	Other	Intergeneic
486	3	179993335	cg22289155	Other	Intergeneic
487	12	46950853	cg23677778	Other	Intergeneic
488	6	27243037	cg21643086	Other	Intergeneic
489	15	47565787	cg10477878	Other	Intergeneic
490	13	87224101	cg21275368	Other	Intergeneic
491	19	37892209	cg26361327	Other	Intergeneic
492	18	5888060	cg06977186	Other	Intergeneic
493	17	39010551	cg07958689	Other	Intergeneic
494	6	24975742	ch.6.25083721F	Other	Intergeneic
495	2	14336631	cg12676991	Other	Intergeneic
496	6	26330589	cg13569146	Other	Intergeneic
497	11	34608061	cg22314759	Other	Intergeneic
498	5	3779072	cg18924848	Other	Intergeneic
499	16	65800351	ch.16.1425090F	Other	Intergeneic
500	5	180097910	cg23214352	Other	Intergeneic
501	14	63131938	cg20468787	Other	Intergeneic
502	7	112135961	cg23222472	Other	Intergeneic
503	10	86910486	cg19088503	Other	Intergeneic
504	6	94550257	ch.6.94606978F	Other	Intergeneic
505	14	88608773	cg13997469	Other	Intergeneic
506	7	125664997	ch.7.125452233F	Other	Intergeneic
507	3	151555029	cg21509105	Other	Intergeneic
508	1	2689171	cg21584800	Other	Intergeneic
509	7	69058543	cg12999084	Other	Intergeneic

510	2	2757161	cg18663259	Other	Intergeneic
511	5	3187456	cg05712938	Other	Intergeneic
512	15	35405377	ch.15.33192669F	Other	Intergeneic
513	7	51658747	cg12988117	Other	Intergeneic
514	19	42439146	cg16700658	Other	Intergeneic
515	2	173575577	cg08641935	Other	Intergeneic
516	2	124194392	ch.2.123910862R	Other	Intergeneic
517	2	227291401	cg25101764	Other	Intergeneic
518	5	56790874	cg06821992	Other	Intergeneic
519	6	164614646	cg10413861	Other	Intergeneic
520	5	1742623	cg12501402	Other	Intergeneic
521	X	124332875	cg09740875	Other	Intergeneic
522	3	59501481	cg06146466	Other	Intergeneic
523	3	177416888	ch.3.3451078R	Other	Intergeneic
524	6	78002283	cg13663057	Other	Intergeneic
525	2	227342994	cg13165983	Other	Intergeneic
526	14	98097986	cg10319905	Other	Intergeneic
527	11	49073835	cg06445586	Other	Intergeneic
528	17	32633974	cg12243622	Other	Intergeneic
529	13	55812567	ch.13.54710568F	Other	Intergeneic
530	3	16779726	cg27614376	Other	Intergeneic
531	12	97951712	cg27109238	Other	Intergeneic
532	16	52641824	cg10109421	Other	Intergeneic
533	13	91343397	ch.13.90141398F	Other	Intergeneic
534	13	35200731	ch.13.381084F	Other	Intergeneic
535	10	62576479	cg09868354	Other	Intergeneic
536	X	138525808	cg09148853	Other	Intergeneic
537	3	8041501	cg16361249	Other	Intergeneic
538	10	93412294	cg16377790	Other	Intergeneic
539	1	88167694	ch.1.87940282F	Other	Intergeneic

540	21	25582901	cg25946965	Other	Intergeneic
541	7	136322567	cg22918741	Other	Intergeneic
542	10	54203646	cg23803709	Other	Intergeneic
543	7	77269758	ch.7.1700983F	Other	Intergeneic
544	5	72750474	cg14102740	Other	Intergeneic
545	13	31972778	ch.13.315182R	Other	Intergeneic
546	11	34608041	cg16935203	Other	Intergeneic
547	4	24274965	cg15650745	Other	Intergeneic
548	8	145910623	cg24497813	Other	Intergeneic
549	X	136921659	cg17550929	Other	Intergeneic
550	21	30188220	ch.21.284298R	Other	Intergeneic
551	6	67606015	ch.6.1433484F	Other	Intergeneic
552	16	51475605	cg27268835	Other	Intergeneic
553	14	86687661	cg07450805	Other	Intergeneic
554	5	158534530	cg19070856	Other	Intergeneic
555	4	19777808	cg24770408	Other	Intergeneic
556	1	171407789	cg25061682	Other	Intergeneic
557	3	44895117	ch.3.44870121R	Other	Intergeneic
558	12	116784790	cg16326611	Other	Intergeneic
559	13	56790593	cg23817132	Other	Intergeneic
560	14	63112650	cg11159234	Other	Intergeneic
561	3	70048377	cg06341047	Other	Intergeneic
562	2	30146945	cg07060894	Other	Intergeneic
563	17	68663564	ch.17.66175159R	Other	Intergeneic
564	10	10462173	ch.10.290763R	Other	Intergeneic
565	2	166937869	cg13875008	Other	Intergeneic
566	16	3201981	cg06643150	Other	Intergeneic
567	17	37309414	cg10213328	Other	Intergeneic
568	12	65174660	cg14078059	Other	Intergeneic
569	3	75862225	ch.3.1652793F	Other	Intergeneic

570	3	19740602	cg07849811	Other	Intergeneic
571	X	150864703	cg09026179	Other	Intergeneic
572	4	24043352	cg25659893	Other	Intergeneic
573	1	214435182	cg16682225	Other	Intergeneic
574	11	112693829	cg24986840	Other	Intergeneic
575	2	5689033	cg13362028	Other	Intergeneic
576	6	166270367	cg16206344	Other	Intergeneic
577	15	26328941	cg23527974	Other	Intergeneic
578	20	52793130	cg25305530	Other	Intergeneic
579	3	88930955	ch.3.89013645F	Other	Intergeneic
580	10	32254360	ch.10.819441R	Other	Intergeneic
581	2	8416612	ch.2.224494R	Other	Intergeneic
582	13	47994694	cg18423626	Other	Intergeneic
583	11	23248542	ch.11.535384R	Other	Intergeneic
584	1	89027289	ch.1.88799877F	Other	Intergeneic
585	3	117075024	cg12837919	Other	Intergeneic
586	19	12904162	cg15317793	Other	Intergeneic
587	3	175640408	ch.3.3414728R	Other	Intergeneic
588	X	141126129	ch.X.2058079F	Other	Intergeneic
589	4	125375437	ch.4.125594887R	Other	Intergeneic
590	X	34404521	cg17001761	Other	Intergeneic
591	12	66119400	cg18758900	Other	Intergeneic
592	2	103593390	cg15128147	Other	Intergeneic
593	1	14477617	cg23803120	Other	Intergeneic
594	6	14426636	ch.6.14534615F	Other	Intergeneic
595	2	21874939	cg16784006	Other	Intergeneic
596	2	240866924	cg20598190	Other	Intergeneic
597	6	32774788	cg22862357	Other	Intergeneic
598	5	36450470	ch.5.36486227R	Other	Intergeneic
599	12	73586034	ch.12.71872301F	Other	Intergeneic

600	12	49046495	ch.12.973812R	Other	Intergeneic
601	12	45356182	ch.12.902977F	Other	Intergeneic
602	10	44782092	cg10211193	Other	Intergeneic
603	11	86529655	cg11873113	Other	Intergeneic
604	14	22385791	cg27268120	Other	Intergeneic
605	3	116996975	cg10462597	Other	Intergeneic
606	14	78838275	ch.14.1202858F	Other	Intergeneic
607	6	127741813	cg12214090	Other	Intergeneic
608	16	49318747	cg26786800	Other	Intergeneic
609	5	53172915	ch.5.53208672R	Other	Intergeneic
610	6	139690828	cg18472160	Other	Intergeneic
611	11	22488445	cg17344099	Other	Intergeneic
612	7	142421812	cg19735804	Other	Intergeneic
613	5	1316636	cg10441424	Other	Intergeneic
614	11	121298154	cg13683424	Other	Intergeneic
615	X	16490374	ch.X.16400295F	Other	Intergeneic
616	8	130738823	cg17837330	Other	Intergeneic
617	13	54740236	cg15157312	Other	Intergeneic
618	6	29621467	cg14193550	Other	Intergeneic
619	4	43876927	cg21164813	Other	Intergeneic
620	5	12865512	cg16107470	Other	Intergeneic
621	10	120437770	ch.10.2535095F	Other	Intergeneic
622	16	65794720	cg10129884	Other	Intergeneic
623	8	40059435	cg20660197	Other	Intergeneic
624	2	70530862	cg17962671	Other	Intergeneic
625	5	6159585	cg09440150	Other	Intergeneic
626	10	132003854	cg13800652	Other	Intergeneic
627	8	49891730	cg08374859	Other	Intergeneic
628	14	35135441	cg23429457	Other	Intergeneic
629	13	58655819	cg09034331	Other	Intergeneic

630	5	97834710	ch.5.97862610R	Other	Intergeneic
631	7	23518135	cg10437900	Other	Intergeneic
632	1	18238511	ch.1.620704R	Other	Intergeneic
633	1	174947362	ch.1.173213985R	Other	Intergeneic
634	5	164258317	ch.5.3099968F	Other	Intergeneic
635	12	112849407	cg22190774	Other	Intergeneic
636	15	92065606	cg18774857	Other	Intergeneic
637	4	101220412	cg27216899	Other	Intergeneic
638	13	39068618	cg22937571	Other	Intergeneic

These CpGs were identified DNA methylation data from Yang *et al* [15]. The base position has been based on National Center for Biotechnology Information genome build 37.

Supplementary Table 2: Differential methylation analysis of 109 identified CpGs in the TCGA colon cancer data

SN	Chromosome	Base position	CpG	Gene	Normol (μ)	Cancer(μ)	P-value
1	1	152087267	cg22603037	TCHH	0.86	0.54	$6.32x10^{-17}$
2	1	44287964	cg23290313	ST3GAL3	0.79	0.53	$6.31x10^{-14}$
3	1	234040045	cg10878114	SLC35F3	0.44	0.67	$2.23x10^{-13}$
4	1	151967449	cg06698332	S100A10	0.5	0.44	0.02
5	1	1981816	cg22865720	PRKCZ	0.01	0.02	0.08
6	1	114503218	ch.1.2681285F	HIPK1	0.17	0.19	0.11
7	1	21074008	ch.1.705736F	HP1BP3	0.1	0.12	0.14
8	2	165998136	cg16631432	SCN3A	0.76	0.46	$7.40x10^{-15}$
9	2	192543258	cg15794798	OBFC2A	0.11	0.14	$2.12x10^{-6}$
10	2	100175805	cg17165836	AFF3	0.12	0.13	0.87
11	2	153575717	cg07491444	ARL6IP6	0.07	0.08	0.92
12	3	159590447	cg09811510	SCHIP1	0.66	0.31	$7.54x10^{-17}$
13	3	35730993	cg15459537	ARPP-21	0.52	0.35	$3.73x10^{-10}$
14	3	25635650	cg07405178	RARB	0.15	0.19	$1.06x10^{-4}$
15	3	15058168	ch.3.343413R	NR2C2	0.07	0.1	$2.94x10^{-4}$
16	3	194136354	ch.3.3822654R	ATP13A3	0.11	0.11	0.2
17	3	73591147	cg09236445	PDZRN3	0.67	0.63	0.41
18	3	52719268	cg09817993	GNL3	0.09	0.1	0.87
19	4	72635202	cg24806812	GC	0.87	0.59	$1.96x10^{-16}$
20	4	85766242	ch.4.1647744F	WDFY3	0.07	0.09	$1.18x10^{-3}$
21	4	121668750	ch.4.2245532F	PRDM5	0.07	0.09	0.02
22	4	25235765	cg12931625	PI4K2B	0.05	0.05	0.06
23	4	110553306	ch.4.2065340F	CCDC109B	0.08	0.1	0.09
24	4	184365198	cg24787081	CDKN2AIP	0.07	0.06	0.11
25	4	47033180	cg21472546	GABRB1	0.55	0.57	0.26
26	4	154707153	cg11467638	SFRP2	0.12	0.13	0.31
27	4	155471778	cg11404039	PLRG1	0.09	0.1	0.41
28	5	167028535	cg15651267	ODZ2	0.83	0.43	$1.75x10^{-19}$

29	5	155909246	cg24132325	SGCD	0.82	0.44	$2.28x10^{-18}$
30	5	113805552	cg25486361	KCNN2	0.81	0.69	$1.24x10^{-6}$
31	5	158634905	cg26362852	RNF145	0.12	0.15	$2.06x10^{-3}$
32	5	135701422	cg17275074	TRPC7	0.24	0.23	0.01
33	5	78281964	cg26802063	ARSB	0.15	0.18	0.06
34	5	156097036	cg15160274	SGCD	0.9	0.89	0.3
35	5	159846543	cg15333689	SLU7	0.05	0.07	0.41
36	5	171653553	ch.5.3268483F	UBTD2	0.13	0.14	0.87
37	6	30655567	cg23903723	KIAA1949	0.23	0.22	0.04
38	6	167200499	cg18495191	RPS6KA2	0.31	0.29	0.27
39	6	1836850	cg21478123	GMDS	0.8	0.8	0.28
40	6	160211006	cg07151830	TCP1	0.06	0.06	0.41
41	7	48319696	cg10626169	ABCA13	0.8	0.47	$6.31x10^{-14}$
42	7	107221074	cg11785538	BCAP29	0.29	0.23	$2.14x10^{-5}$
43	7	150924351	cg10436877	ABCF2	0.22	0.24	0.03
44	7	79765394	cg25702790	GNAI1	0.08	0.1	0.92
45	8	12974556	cg06103928	DLC1	0.83	0.74	$7.16x10^{-5}$
46	8	38174205	ch.8.903080R	WHSC1L1	0.1	0.11	0.08
47	9	113800909	cg06148685	LPAR1	0.05	0.08	0.09
48	10	108674143	cg23024358	SORCS1	0.82	0.65	$8.67x10^{-16}$
49	10	44287023	cg17952824	HNRNPA3P1	0.51	0.38	$2.92x10^{-6}$
50	10	96306185	cg10069677	HELLS	0.11	0.14	$1.03x10^{-3}$
51	10	79793495	cg23620279	RPS24	0.09	0.1	0.05
52	10	63661280	cg16253809	ARID5B	0.1	0.11	0.07
53	10	13628544	cg15881990	PRPF18	0.14	0.15	0.15
54	10	13628544	cg15881990	PRPF18	0.14	0.15	0.15
55	10	69913749	cg18986048	MYPN	0.21	0.22	0.83
56	11	128458153	cg16792062	ETS1	0.87	0.51	$2.28x10^{-19}$
57	11	55796572	cg17060964	OR5AS1	0.66	0.29	$7.78x10^{-19}$
58	11	51413644	cg23935054	OR4A5	0.21	0.13	$2.21x10^{-13}$

59	11	63140804	cg10116443	SLC22A9	0.9	0.73	$1.66x10^{-6}$
60	11	108092818	cg12019961	ATM	0.12	0.15	$8.54x10^{-4}$
61	11	55606710	cg12962308	OR5D16	0.93	0.86	0.02
62	11	64795449	cg21821990	SNX15	0.1	0.11	0.59
63	12	23998997	cg06764736	SOX5	0.93	0.84	$4.67x10^{-7}$
64	12	122065180	cg24082347	ORAI1	0.02	0.02	0.05
65	12	113773298	cg12583184	SLC24A6	0.2	0.21	0.16
66	12	67662516	cg18750937	CAND1	0.29	0.29	0.28
67	12	45034784	ch.12.897509F	NELL2	0.11	0.11	0.41
68	12	50527085	ch.12.1019410F	LASS5	0.11	0.12	0.41
69	12	50450889	cg23126949	ACCN2	0.09	0.1	0.45
70	12	58087540	cg18202167	OS9	0.16	0.17	0.69
71	12	50635579	ch.12.1023240F	LIMA1	0.1	0.11	0.73
72	13	60543691	ch.13.865492R	DIAPH3	0.13	0.12	0.34
73	13	49684397	cg17091793	FNDC3A	0.05	0.05	0.41
74	13	28194831	cg07375367	POLR1D	0.19	0.18	0.53
75	14	78869352	cg10828316	NRXN3	0.81	0.43	$1.32x10^{-19}$
76	14	80324276	cg19753609	NRXN3	0.77	0.52	$3.90\ x10^{-12}$
77	14	63508497	cg25609301	KCNH5	0.39	0.37	0.18
78	14	33826344	cg15454195	NPAS3	0.61	0.57	0.26
79	14	24701654	cg07519822	GMPR2	0.05	0.05	0.77
80	15	48515109	cg23530596	SLC12A1	0.62	0.27	$1.32x10^{-19}$
81	15	61346347	cg08099431	RORA	0.9	0.72	$1.50x10^{-10}$
82	15	70390363	cg07020846	TLE3	0.08	0.09	$5.49x10^{-4}$
83	15	99789637	cg25385940	TTC23	0.21	0.19	0.15
84	16	9010914	ch.16.350833F	USP7	0.18	0.17	0.41
85	17	49281558	ch.17.1348593F	MBTD1	0.15	0.16	0.41
86	19	53619086	cg22834281	ZNF415	0.92	0.64	$1.13x10^{-13}$
87	19	20011538	cg27379065	ZNF93	0.09	0.15	$3.56x10^{-5}$
88	19	58790298	cg23548487	ZNF8	0.04	0.08	$3.02x10^{-5}$

89	19	12277357	cg13689563	ZNF136	0.9	0.9	0.38
90	19	4066818	cg10561472	ZBTB7A	0.07	0.07	0.41
91	20	45937282	ch.20.1002962F	ZMYND8	0.22	0.15	$8.00x10^{-8}$
92	20	10026325	ch.20.221631R	ANKRD5	0.12	0.11	0.33
93	22	46114168	ch.22.909671F	ATXN10	0.11	0.14	$8.96x10^{-4}$
94	22	20004611	cg12912949	ARVCF	0.02	0.02	0.18
95	X	102629870	cg13486082	NGFRAP1	0.9	0.78	$1.99x10^{-6}$
96	X	2984799	cg17012513	ARSF	0.24	0.21	$7.73x10^{-4}$
97	X	131547697	cg14520512	MBNL3	0.68	0.59	$8.96x10^{-4}$
98	X	102629912	cg25198830	NGFRAP1	0.87	0.77	$1.60x10^{-3}$
99	X	12993075	cg23376554	TMSB4X	0.19	0.2	0.07
100	X	131547702	cg13633856	MBNL3	0.78	0.76	0.09
101	X	128657727	cg18959966	SMARCA1	0.28	0.36	0.15
102	X	102611412	cg13208102	WBP5	0.25	0.22	0.25
103	X	102611415	cg27464574	WBP5	0.22	0.2	0.27
104	X	153169465	cg20664654	AVPR2	0.53	0.57	0.28
105	X	146312384	cg27167381	MIR506	0.89	0.82	0.34
106	X	77154874	cg10646076	COX7B	0.32	0.29	0.48
107	X	77154996	cg15830530	COX7B	0.23	0.21	0.53
108	X	99892000	cg11509733	TSPAN6	0.14	0.14	0.56
109	X	10087726	cg25497053	WWC3	0.08	0.08	0.92

Supplementary Table 3: Differential expression analysis of identified probes in TCGA colon cancer data

SN	Gene	log2(fold)	P-value
1	ABCF2	-1.12	$3.25x10^{-13}$
2	ACCN2	-1.09	$2.17x10^{-9}$
3	ARID5B	-1.28	$1.35x10^{-3}$
4	ARSB	1.06	0.27
5	ARVCF	-1.05	$1.22x10^{-8}$
6	ATM	1.32	0.41
7	AVPR2	-1.05	$3.00x10^{-7}$
8	BCAP29	1.03	$1.76x10^{-3}$
9	CAND1	-1.18	$1.51x10^{-7}$
10	CDKN2AIP	-1.59	0.06
11	COX7B	-1.25	$9.78x10^{-3}$
12	ETS1	1.04	0.01
13	FNDC3A	-1.15	0.06
14	GABRB1	-1.06	0.02
15	GNL3	-1.10	$1.63x10^{-13}$
16	HNRNPA3P1	-1.09	$9.84x10^{-9}$
17	KIAA1949	-1.41	0.33
18	LPAR1	-1.07	$4.48x10^{-14}$
19	NGFRAP1	-1.16	0.43
20	NMT2	-1.09	$6.92x10^{-9}$
21	NRXN3	-1.05	0.03
22	NTM	-1.31	0.02
23	OS9	-1.13	$5.76x10^{-3}$
24	PIGA	-1.12	$5.85x10^{-7}$
25	PLRG1	-1.35	0.01
26	POLR1D	-1.08	$3.25x10^{-13}$
27	PRKCZ	-1.30	0.06
28	PRPF18	-1.32	0.06
29	RNF145	-1.09	0.19

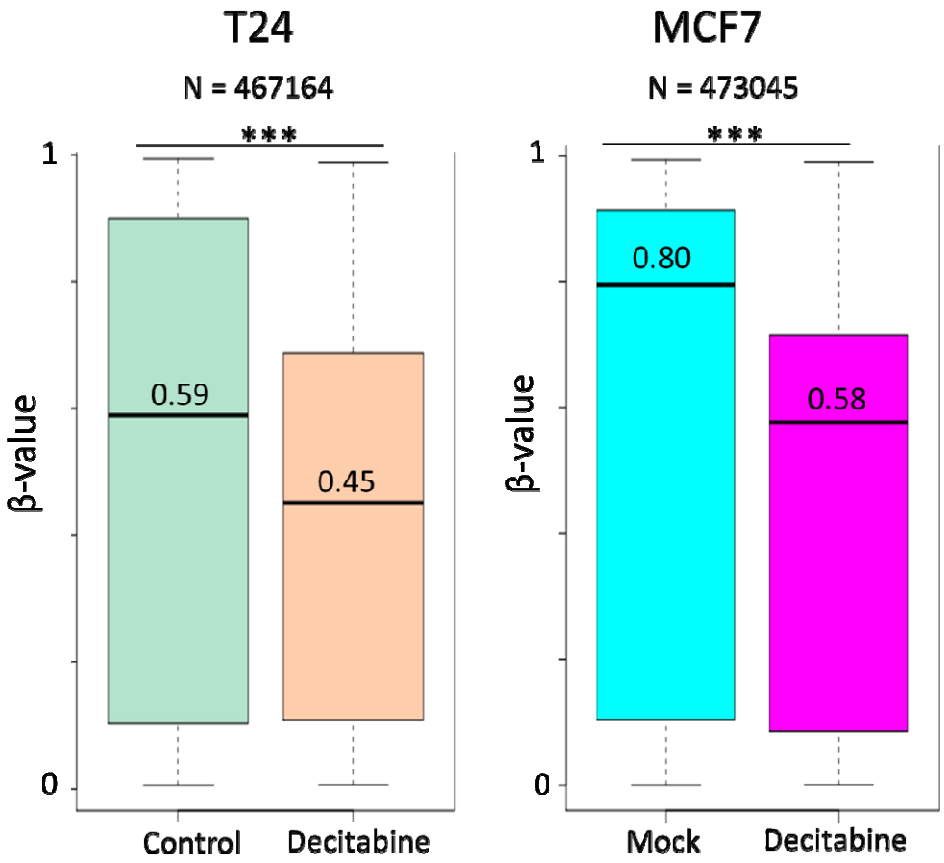
30	RPS24	-1.05	$1.51x10^{-7}$
31	S100A10	-1.07	0.01
32	SEPSECS	-1.21	0.19
33	SLC24A6	-1.22	$1.10x10^{-6}$
34	SLC35F3	-1.05	$1.81x10^{-7}$
35	SLU7	-1.88	0.29
36	SMARCA1	-1.04	$6.62x10^{-3}$
37	TCEAL8	1.00	0.76
38	TCHH	-1.09	0.06
39	TCP1	-1.14	$2.32x10^{-7}$
40	TLE3	1.06	$7.74x10^{-3}$
41	TP73	-1.14	$4.28x10^{-11}$
42	TSPAN6	-1.13	$8.22x10^{-3}$
43	ZBTB7A	-1.13	$4.25x10^{-7}$
44	ZNF8	-1.05	$3.36x10^{-7}$
45	ZNF93	-1.02	$2.40x10^{-3}$
46	ABCA13	1.13	0.59
47	AFF3	-1.09	$3.25x10^{-13}$
48	AGTPBP1	-1.07	$1.99x10^{-9}$
49	ANKRD5	-1.10	$2.32x10^{-7}$
50	ARL6IP6	-1.18	$1.07x10^{-9}$
51	ATP13A3	-1.12	$2.17x10^{-9}$
52	ATXN10	-1.19	$4.74x10^{-4}$
53	CADM1	-1.03	$4.45x10^{-6}$
54	CCDC109B	-1.35	$6.20x10^{-5}$
55	COX7B1	-1.25	$9.78x10^{-3}$
56	CUL3	-1.07	0.46
57	DIAPH3	-1.11	$1.76x10^{-13}$
58	DLC1	-1.01	$7.01x10^{-4}$
59	FAT3	-1.03	0.06

60	GAB2	-1.11	$3.08x10^{-3}$
61	GMDS	-1.04	0.02
62	GMPR2	-1.02	0.08
63	GNAI1	-1.08	$1.32x10^{-9}$
64	HELLS	-1.11	$2.49x10^{-13}$
65	HIPK1	1.00	0.10
66	HP1BP3	-1.19	0.82
67	KCNN2	1.27	0.34
68	LARS2	-1.12	$6.61x10^{-5}$
69	LASS5	-1.09	$8.59x10^{-9}$
70	LIMA1	-1.10	$1.11x10^{-10}$
71	MAP7	-1.11	$6.43x10^{-7}$
72	MBNL3	-1.06	$2.26x10^{-5}$
73	MBTD1	-1.12	$1.76x10^{-3}$
74	MRPS9	1.05	0.04
75	MYPN	-1.12	$8.16x10^{-10}$
76	NELL2	-1.11	$4.36x10^{-7}$
77	NPAS3	-1.08	$1.90x10^{-10}$
78	NR2C2	-1.27	$9.78x10^{-3}$
79	NRXN31	-1.05	0.03
80	OBFC2A	-1.03	$3.08x10^{-3}$
81	ODZ2	-1.04	$1.50x10^{-6}$
82	ORAI1	-1.09	$1.15x10^{-4}$
83	PDE1C	-1.10	$6.74x10^{-12}$
84	PDZRN3	-1.09	$5.34x10^{-8}$
85	PI4K2B	-1.16	$5.32x10^{-7}$
86	PRDM5	-1.12	0.90
87	RARB	1.11	0.24
88	RFC3	-1.10	$8.98x10^{-14}$
89	RGMB	-1.13	$8.22x10^{-3}$

90	RORA	-1.03	$7.16x10^{-5}$
91	RPS6KA2	-1.17	0.24
92	SCHIP1	-1.04	$1.70x10^{-4}$
93	SCN3A	-1.08	$1.76x10^{-13}$
94	SFRP2	1.00	$3.39x10^{-5}$
95	SGCD	1.01	0.01
96	SMYD4	-1.08	$5.76x10^{-3}$
97	SNX15	-1.10	$1.42x10^{-4}$
98	SORCS1	-1.08	$3.39x10^{-13}$
99	SOX5	-1.10	$2.32x10^{-7}$
100	ST3GAL3	-1.09	$1.04x10^{-7}$
101	TRNAU1AP	1.14	0.52
102	TTC23	-1.15	$2.74x10^{-6}$
103	UBE2L3	-22.63	0.45
104	UBTD2	-1.06	$2.49x10^{-3}$
105	USP7	-1.07	$2.32x10^{-4}$
106	WBP5	-1.21	0.45
107	WDFY3	1.14	0.52
108	WHSC1L1	-1.22	0.07
109	WWC3	-1.09	$3.52x10^{-5}$
110	ZMYND8	-1.09	$2.90x10^{-6}$
111	ZNF136	3.17	0.59
112	ZNF415	-1.09	$5.58x10^{-9}$

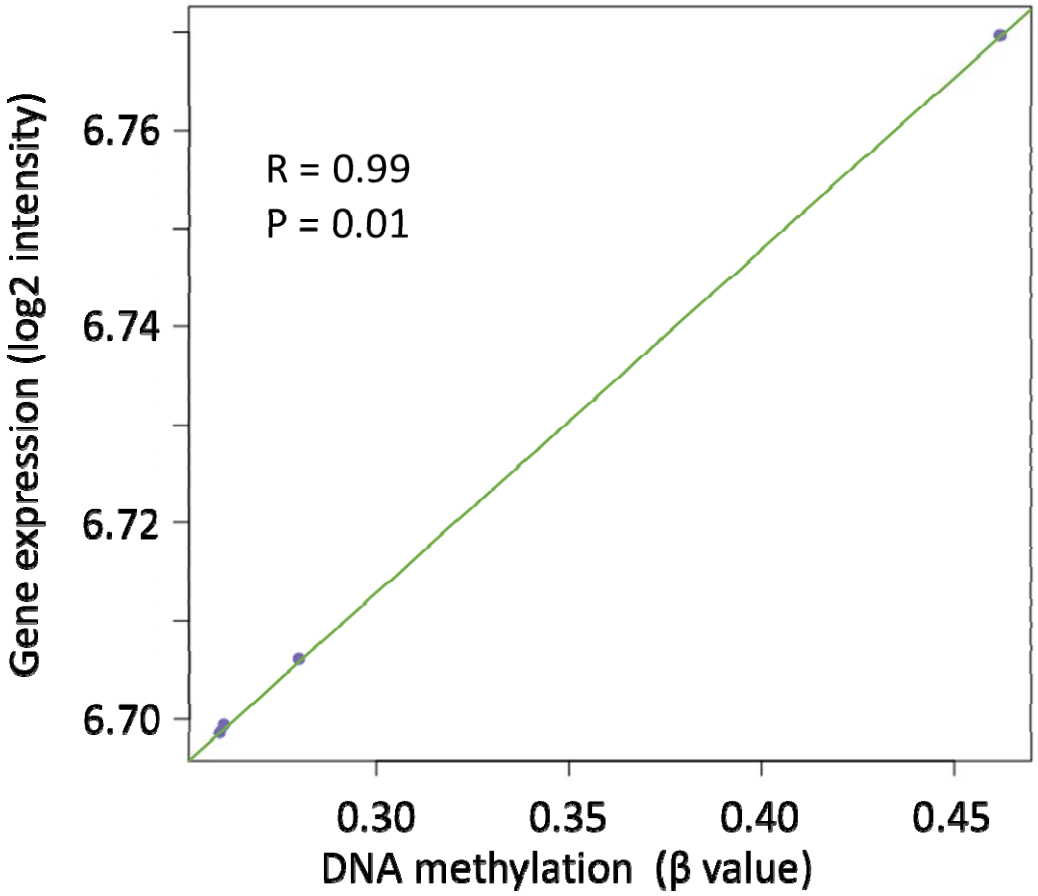
517

Supplementary figure 1: Box plot showing a decrease in methylation level of loci other than identified CpGs in T24 and MCF7 cell line after decitabine treatment.

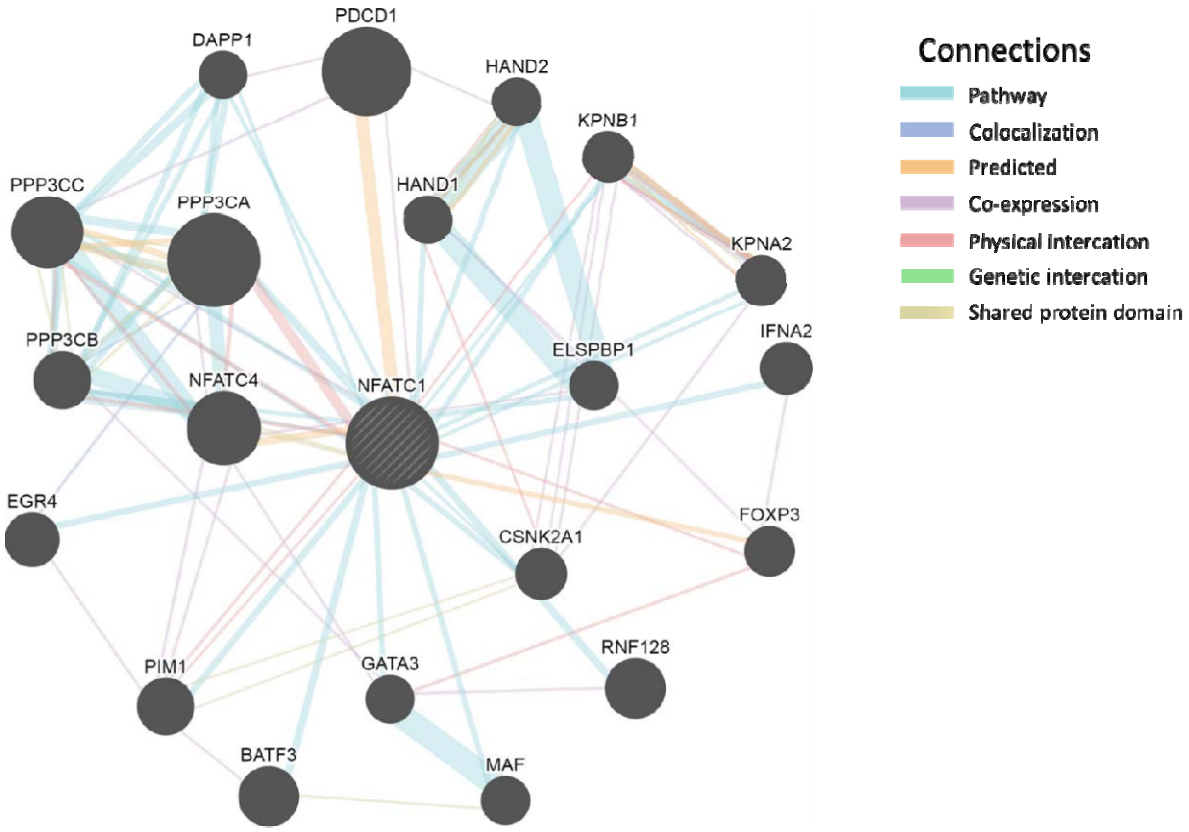


N denotes the total number of CpGs analyzed in the data. The data from the study by Han *et al* [16] (GSE41525) and Leadem *et al* [17] (GSE97483) has been shown for T24 and MCF7 cells respectively

Supplementary figure 2: Scatter plot showing the correlation between RORA expression level and methylation level at CpGs cg08099431 in HCT116 cell line.

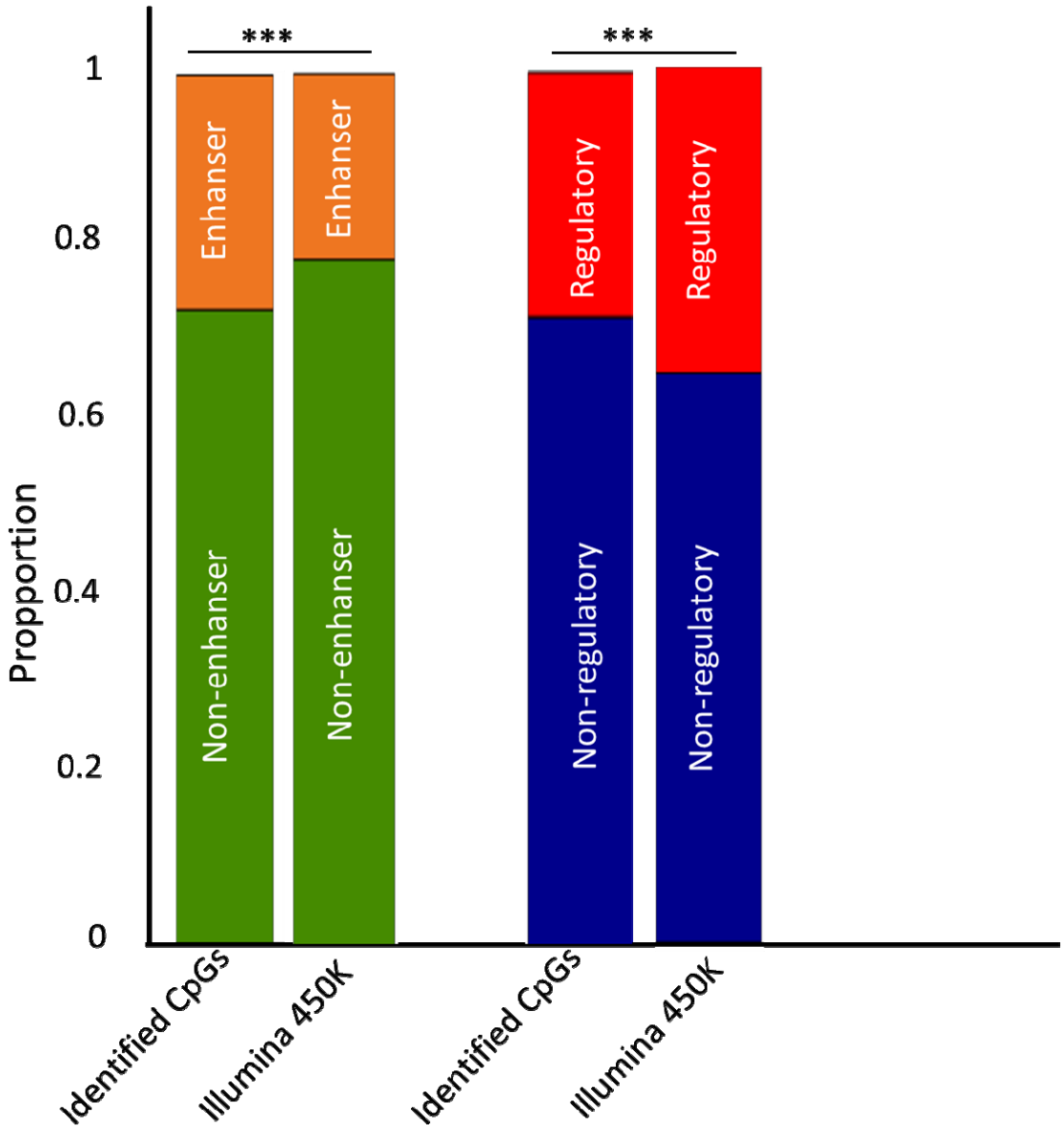


Supplementary figure 3: Interaction network of genes regulated by NFATc1.



Only interaction among those genes that are directly connected to NFATc1 has been shown using network construction from GENEMANIA. The size of the gene nodes is proportional to gene score calculated by GENEMANIA using label propagation algorithm that indicates the relevance of each gene to the original list based on the selected networks.

Supplementary figure 4: Enrichment analysis of identified CpGs in enhancer and regulatory region



Eenrichment of identified CpGs among regulatory region of the genome. CpGs with increased methylations were enriched in enhancer region (21% of identified CpGs were in enhancer region as compared to 26% of total CpGs present in 450K chip, P = $3.36x10^{-4}$) and were depleted in other regulatory regions such as promoter and non-promoter associated cell type-specific or general regulatory regions represented by transcription factor binding sites and DNA hypersensitivity elements (25% of identified CpGs were in other regulatory region as compared to 28% of total

*CpGs in 450K beadchip, P = 2.65x10^{-4}). The proportions of identified CpGs in enhancer and regulatory region have been shown as segmented barplot (upper segment). ***P<0.0005*