Biology of Human Tumors

Clinical Cancer Research

Kidney Cancer Is Characterized by Aberrant Methylation of Tissue-Specific Enhancers That Are Prognostic for Overall Survival

Caroline Y. Hu[1], Davoud Mohtat[1], Yiting Yu[1], Yi-An Ko[2], Niraj Shenoy[1], Sanchari Bhattacharya[1], Maria C. Izquierdo[2], Ae Seo Deok Park[2], Orsolya Giricz[1], Nishanth Vallumsetla[1], Krishna Gundabolu[1], Kristin Ware[3], Tushar D. Bhagat[1], Masako Suzuki[1], James Pullman[3], X. Shirley Liu[4], John M. Greally[1], Katalin Susztak[2], and Amit Verma[1]

Abstract

Purpose: Even though recent studies have shown that genetic changes at enhancers can influence carcinogenesis, most methylomic studies have focused on changes at promoters. We used renal cell carcinoma (RCC), an incurable malignancy associated with mutations in epigenetic regulators, as a model to study genome-wide patterns of DNA methylation at a high resolution.

Experimental Design: Analysis of cytosine methylation status of 1.3 million CpGs was determined by the HELP assay in RCC and healthy microdissected renal tubular controls.

Results: We observed that the RCC samples were characterized by widespread hypermethylation that preferentially affected gene bodies. Aberrant methylation was particularly enriched in kidney-specific enhancer regions associated with H3K4Me1 marks. Various important underexpressed genes, such as SMAD6, were associated with aberrantly methylated, intronic enhancers, and these changes were validated in an independent cohort. MOTIF analysis of aberrantly hypermethylated regions revealed enrichment for binding sites of AP2a, AHR, HAIRY, ARNT, and HIF1 transcription factors, reflecting contributions of dysregulated hypoxia signaling pathways in RCC. The functional importance of this aberrant hypermethylation was demonstrated by selective sensitivity of RCC cells to low levels of decitabine. Most importantly, methylation of enhancers was predictive of adverse prognosis in 405 cases of RCC in multivariate analysis. In addition, parallel copy-number analysis from MspI representations demonstrated novel copy-number variations that were validated in an independent cohort of patients.

Conclusions: Our study is the first high-resolution methylome analysis of RCC, demonstrates that many kidney-specific enhancers are targeted by aberrant hypermethylation, and reveals the prognostic importance of these epigenetic changes in an independent cohort. *Clin Cancer Res; 20(16); 4349–60. ©2014 AACR.*

Introduction

Patterns of DNA methylation are altered in carcinogenesis and play important roles in regulating gene transcription and genomic stability (1). Even though most of the previous studies focused on epigenetic changes at promoters, recent high-resolution studies have revealed that aberrant methylation can affect gene bodies (2). Intragenic methylation has been correlated with changes in gene transcription (3), but it has not been shown clearly whether aberrant intronic methylation affects any regulatory regions of the genome. Recent data have also revealed that enhancers play important roles in regulating gene transcription and their alterations can play roles in carcinogenesis (4–6). These data prompted us to examine the role of aberrant intragenic methylation in cancer using renal cancer as a model and to analyze whether it has any clinical implications in this incurable disease.

Renal cell carcinoma (RCC) affects more than 200,000 individuals worldwide and is the ninth most common cancer in the United States with a rising incidence (7). The treatment for RCC confined to the parenchyma is primary surgical and has an overall survival (OS) of 60% to 70%. However, advanced RCC carries a very poor prognosis with limited therapeutic options (8). RCC comprises a

Authors' Affiliations: [1]Albert Einstein College of Medicine, Bronx, New York. [2]Renal Electrolyte and Hypertension Division, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania. [3]Department of Pathology, Montefiore Medical Center, New York, New York. [4]Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, Massachusetts

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

C.Y. Hu, D. Mohtat, and Y. Yu contributed equally to this article.

Corresponding Authors: Amit Verma, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Chanin 302B, Bronx, NY 10461. Phone: 718-430-8761; Fax: 718-430-8702; E-mail: amit.verma@einstein.yu.edu; and Katalin Susztak, ksusztak@mail.med.upenn.edu

doi: 10.1158/1078-0432.CCR-14-0494

Translational Relevance

Even though high-resolution studies have shown that changes in DNA methylation can affect intragenic (gene body) regions of the genome, the clinical and biologic significance of these changes is not clearly defined. We demonstrate that kidney cancer is characterized by increased amounts of methylation that affects introns. These areas of increased methylation preferentially affect regulatory regions called enhancers. Importantly, increased methylation at enhancers is predictive of worse clinical outcomes in a large independent dataset in a multivariate analysis, demonstrating the importance of these epigenetic changes.

multitude of histologic subtypes, each with a different clinical phenotype and genetic abnormality. Clear cell subtype is the most common and has a high incidence of alterations on chromosome 3 and in the VHL (Von Hippel–Lindau) gene (7). The VHL/HIF pathway has been shown to play an important role in RCC and cases can be subgrouped on the basis of their VHL and HIF expression (9). RCC is resistant to radiotherapy and chemotherapy, and approved kinase inhibitors have led to only minimal improvements in OS (10). Recent genetic studies also indicate mutations of different chromatin modifying enzymes, such as PBRM1, BAP1, SETD2, and KDM5C in RCC (11, 12). These studies suggest that the epigenetic dysregulation occurs in RCC and needs to be studied at high resolution.

Several experimental approaches are available to determine genome-wide DNA methylation levels. Most of these techniques are based on restriction enzyme digestion or DNA immuneprecipitation with antibodies that bind to methylated CpGs (13). The HELP (HpaII tiny fragment enrichment by ligation-mediated PCR) assay relies on differential digestion by a pair of enzymes, HpaII and MspI, which differ on the basis of their methylation sensitivity. The HpaII and MspI genomic representations can be cohybridized to a custom microarray and their ratio used to indicate the methylation of particular CCGG sites at these loci. The HELP assay has been shown to be a robust discovery tool and has been successful in revealing novel epigenetic alterations in leukemias, myelodyplasia, and esophageal cancer (14–16). Most studies on DNA methylation in RCC have been single locus studies and have focused only on promoters and CpG islands (7, 17). Newer data have shown that the non-CpG island loci are very important in gene regulation (18). Furthermore, newer higher-resolution assays reveal that gene body methylation may be even more important in gene regulation than promoter methylation (19). A recent genome-wide study revealed hypermethylation in RCC (20) and further necessitates the study of these changes at higher resolution to examine the role of aberrant gene body methylation in renal cell cancer.

In addition to epigenetic alterations, RCC is also characterized by many cytogenetic abnormalities that may contribute to its pathogenesis. We have developed an integrated genomics and epigenomics platform and used it on RCC samples. Our studies showed that methylation changes could be seen in RCC and affect intronic enhancers. Epigenetic changes at enhancers were also highly prognostic in patients. Also, both novel and well-characterized genomic copy-number changes were also found in these RCC samples. In addition to demonstrating the involvement of novel intronic regulatory elements by aberrant methylation in our and other independent datasets, we also demonstrate the translational potential of DNA methyltransferase inhibitors in RCC.

Materials and Methods

Patient samples, microdissection, and nucleic acid extraction

In total, 13 RCC samples and 13 control samples were collected from tumor nephrectomies performed at the Montefiore Medical Center. Samples were collected under the Institutional Review Board protocol approved by the Albert Einstein College of Medicine (Bronx, NY). Kidney samples were obtained from the living allograft donors, surgical nephrectomies, and left-over portions of diagnostic kidney biopsies. Nephrectomies were anonymized with the corresponding clinical information and were collected by an individual who was not involved in the research protocol. Tissue was placed into RNALater and manually microdissected at 4°C for the tubular compartment (as renal cell cancer originates from tubular epithelial cells; ref. 21). We used a Zeiss steromicroscope under $\times 60$ magnifications for the microdissection. Genomic DNA was isolated using the dialysis tubing method, as performed and described previously (1). RNA was extracted using Qiagen RNeasy Mini Kits. An additional seven nontumor and tumor pairs were used for qRT-PCR–based confirmation studies.

Genome-wide DNA methylation analysis using the HELP assay

The HELP assay was carried out as previously published (22). Intact DNA of high molecular weight was corroborated by electrophoresis on 1% agarose gel in all cases. One microgram of genomic DNA was digested overnight with either HpaII or MspI (NEB). The following day the reactions were extracted once with phenol-chloroform and resuspended in 11 μL of 10 mmol/L Tris-HCl pH 8.0. The digested DNA was used to set up an overnight ligation of the JHpaII adapter using T4 DNA ligase. The adapter-ligated DNA was used to carry out the PCR amplification of the HpaII and MspI-digested DNA as previously described (23). Both amplified fractions were submitted to Roche-NimbleGen, Inc. for labeling and hybridization onto a human hg18 custom-designed oligonucleotide array (50 mers) covering 1.3 million HpaII amplifiable fragments (HAF). HELP microarray data have been submitted to the GEO database for public access (GEO accession: GSE49420; refs. 15, 22).

All microarray hybridizations were subjected to extensive quality control. Uniformity of hybridization was evaluated using a modified version of a previously published algorithm (16) adapted for the NimbleGen platform, and any hybridization with strong regional artifacts was discarded.

Quantitative DNA methylation analysis by MassArray Epityping

Validation of HELP microarray findings was carried out by MALDI–TOF (matrix-assisted laser desorption/ionization–time-of-flight) mass spectrometry using EpiTyper by MassArray (Sequenom) on bisulphite-converted DNA as previously described (24, 25). MassArray primers were designed to cover the flanking HpaII sites for a given HAF, as well as any other HpaII sites found up to 2,000 bp upstream of the downstream site and up to 2,000 bp downstream of the upstream site, to cover all possible alternative sites of digestion.

Gene copy-number analysis using MspI representations from the HELP assay

The MspI representation in a HELP assay is not affected by cytosine methylation but is instead dependent on the amount of DNA available, and, thus, was used to detect copy-number variation (CNV), as described previously (10, 22). VHL gene deletions are seen commonly in previous studies (20) and were found in 3 of 13 cases examined.

qRT-PCR on genomic DNA for copy-number validations. gDNA from kidney samples were isolated using phenol chloroform extraction. qRT-PCR was performed using the SYBR Green method with gene-specific primers with the ABI 7900HT machine. The data were normalized using GAPDH as housekeeping genes and the $\Delta\Delta C_t$ method.

qRT-PCR for SMAD6

cDNA from CAKI kidney cancer cells that were treated with 5-azacytidine (0.5 μmol/L) for 5 days was used for qRT-PCR using the SYBR Green method with gene-specific primers using the ABI 7900HT machine. The data were normalized using hypoxanthine phosphoribosyltransferase as the housekeeping gene and the $\Delta\Delta C_t$ method.

Gene-expression profiling

RNA integrity was corroborated with the Agilent Bioanalyzer 2100. RNA (100 ng/μL; 3 μL) was submitted to the Genomics Facility, Albert Einstein College of Medicine for gene-expression studies using the Human Affymetrix U133 2.0 arrays. RNA was labeled using Affymetrix two-cycle labeling and amplification kit.

HELP data processing and analysis

Signal intensities at each HpaII amplifiable fragment were calculated as a robust (25% trimmed) mean of their component probe-level signal intensities. Any fragments found within the level of background MspI signal intensity, measured as 2.5 mean-absolute-differences above the median of random probe signals, were categorized as "failed." These "failed" loci, therefore, represent the population of fragments that did not amplify by PCR, whatever the biologic (e.g., genomic deletions and other sequence errors) or experimental cause. On the other hand, "methylated" loci were so designated when the level of HpaII signal intensity was similarly indistinguishable from background. PCR-amplifying fragments (those not flagged as either "methylated" or "failed") were normalized using an intra-array quantile approach, wherein HpaII/MspI ratios are aligned across density-dependent sliding windows of fragment size-sorted data. The $\log_2$(HpaII/MspI) was used as a representative for methylation and analyzed as a continuous variable. For most loci, each fragment was categorized as either methylated, if the centered log HpaII/MspI ratio was less than zero, or hypomethylated if on the other hand the log ratio was greater than zero. Data were deposited in NCBI's GEO database (GSE49420).

Microarray data analysis

Unsupervised clustering of HELP data by hierarchical clustering (1-Pearson correlation distance and Ward agglomeration method) was performed using the statistical software R version 2.6.2. A two-sample *t* test was used for each gene to summarize methylation differences between groups. Genes were ranked on the basis of this test statistic and a set of top differentially methylated genes with an observed log fold change of >1 between group means was identified. Genes were further grouped according to the direction of the methylation change (hypomethylated vs. hypermethylated), and the relative frequencies of these changes were computed among the top candidates to explore global methylation patterns. Validations with MassArray showed good correlation with the data generated by the HELP assay. MassArray analysis validated significant quantitative differences in methylation for differentially methylated genes selected by our approach.

Pathway analysis and transcription factor–binding site analysis

Using the Ingenuity Pathway Analysis (IPA) software, we carried out an analysis of the biologic information retrieved by each of the individual platforms. Enrichment of genes associated with specific canonical pathways was determined relative to the Ingenuity knowledge database for each of the individual platforms and the integrated analysis at a significance level of $P < 0.01$. The list of hypermethylated genes was examined for enrichment of conserved gene-associated transcription factor–binding sites using the Cis-regulatory Element Annotation System (CEAS) program (26). The functional gene sets were obtained from the gene ontology (27).

Genomic annotations

Genomic coordinates were obtained from HG18 build of the human genome from the University of California, Santa Cruz (UCSC) browser. Genomic regions 2 kb

upstream and downstream of the transcription start sites were annotated as promoters. Of note, 2-kb flanking regions around the edges of CpG islands were annotated as CpG shores.

Cell lines. Normal human kidney tubular epithelial cell line HK2 and renal cell cancer cell lines 786-O and 769-P were purchased from the ATCC. HKP-8 cells were kindly provided by Lorainne Racusen (Johns Hopkins University, Baltimore, MD). Cells were cultured in DMEM/F12 medium supplemented with 2.5% fetal bovine serum, antibiotics, insulin, transferrin, and selenium.

BrdUrd assay. Cell proliferation was studied using the BrdUrd (bromodeoxyuridine)-based Colorimetric Cell Proliferation Assay Kit from Exalpha Biological Inc. following the manufacturer's instructions. Cells were incubated with 0.5 μmol/L after the seeding and incubated with BrdUrd for 48 hours before harvest. The data are normalized to the baseline (without decitabine) proliferation levels.

Overlap with histone modification studies. All the ChIP-seq data were downloaded from Roadmap database with GEO accession numbers specified below (NCBI Build GRCh37/UCSC Build hg19). The following datasets were used: Adult Kidney BI.27 (GEO accession: GSM670025), Adult Kidney Input 27, AK H3K4me3 BI.27 (GEO accession: GSM621648), AK H3K9ac BI.27 (GEO accession: GSM772811), AK H3K36me3 BI.27 (GEO accession: GSM621634), and AK Input BI.27 (GEO accession: GSM621638).

Integrated study of RCC DMR and histone marks. All the ChIP-seq data were downloaded from Roadmap database with GEO Series accession number of GSE19465. The Genome assembly is NCBI Build GRCh37/UCSC Build hg19. The nine cell-line ChromHMM data were obtained from the UCSC genome browser (28). For adult kidney, ChromHMM annotation map was generated using ChromHMM2 (29).

TCGA data analysis

Data for RNA-seq for RCC samples were obtained from The Cancer Genome Atlas (TCGA) portal. These included 405 RCC samples and 68 nontumor kidney controls (20). For methylation analysis, a mean methylation score for all nontumor controls was calculated. Difference in methylation for each locus for each tumor sample was calculated using the control mean. A methylation difference of more than >20% was considered significant and used to flag the locus as methylated. Cumulative score of number of methylated loci for promoters, gene bodies, and enhancers was used for Kaplan–Meier curves.

Results

Renal cell cancer is characterized by widespread changes in DNA cytosine methylation

We analyzed the methylome of clear cell renal cell cancer (CCRCC) by the HELP assay, to determine the CpG methylation status of 1.3 million loci across the genome. Thirteen cases of histologically verified CCRCC tumor samples were compared with 13 control microdissected proximal tubules from a nontumor part of nephrectomy samples (Supplementary Table S1 describing the clinical and pathologic characteristics of the tumors). Unsupervised hierarchal clustering showed that the controls formed a cluster that was distinct from CCRCC samples, demonstrating epigenetic dissimilarity between these groups (Fig. 1A). The methylation profiling of RCC was not influenced by patient demographics and epigenetic similarity between clusters of samples was independent of age, gender, history of diabetes mellitus and hypertension, or VHL mutational status. We next wanted to determine whether gene-expression patterns differed between RCC and controls. Unsupervised hierarchal clustering showed that the controls formed a cluster that was distinct from RCC samples, demonstrating global gene-expression dissimilarity between these groups (Fig. 1B). These data demonstrate that widespread differences are seen both at the level of the methylome and transcriptome in RCC.

RCC is characterized by genome-wide hypermethylation that affects gene bodies

Having demonstrated epigenetic dissimilarity between RCC and control samples, we next determined the qualitative epigenetic differences between these groups by performing a supervised analysis of the respective DNA methylation profiles. A volcano plot comparing the differences between mean methylation of individual loci in RCC versus control samples plotted against the significance [log (*P* value) based on *t* test] of the difference was used to represent these data in Fig. 2A. We observed that there were significantly increased numbers of hypermethylated loci in RCC ($n = 3{,}378$ hypermethylated vs. 43 hypomethylated loci, *t* test *P* value of <0.0003) when compared with controls (Fig. 2A). This preponderance of hypermethylation was seen with stringent cutoff of a *P* value of < 0.0003 with an FDR <0.1 that was used to determine differentially methylated loci. This striking genome-wide hypermethylation is in contrast with previous reports demonstrating global hypomethylation in other solid tumors (1, 30). The hypermethylation affected all parts of the genome (Fig. 2B–D) and a high proportion of gene bodies (by refseq annotation) were found to be affected by aberrant methylation (Fig. 2E). Interestingly, a greater proportion of CpG shores was also found to be affected by aberrant hypermethylation when compared CpG islands, consistent with other recent observations implicating these genomic regions as targets of aberrant methylation in cancer (18). Correlation with changes in gene expression showed that changes in DNA methylation in tumors were associated significantly with corresponding changes in gene expression in all of these regions, thus demonstrating the significance of these epigenetic changes (Fig. 2F). We also quantitatively estimated the aberrant methylation of selected loci by MALDI–TOF (MassArray; Sequenom) and observed a strong correlation with the findings of our HELP microarrays, demonstrating the validity of our findings (Supplementary Fig. S1).

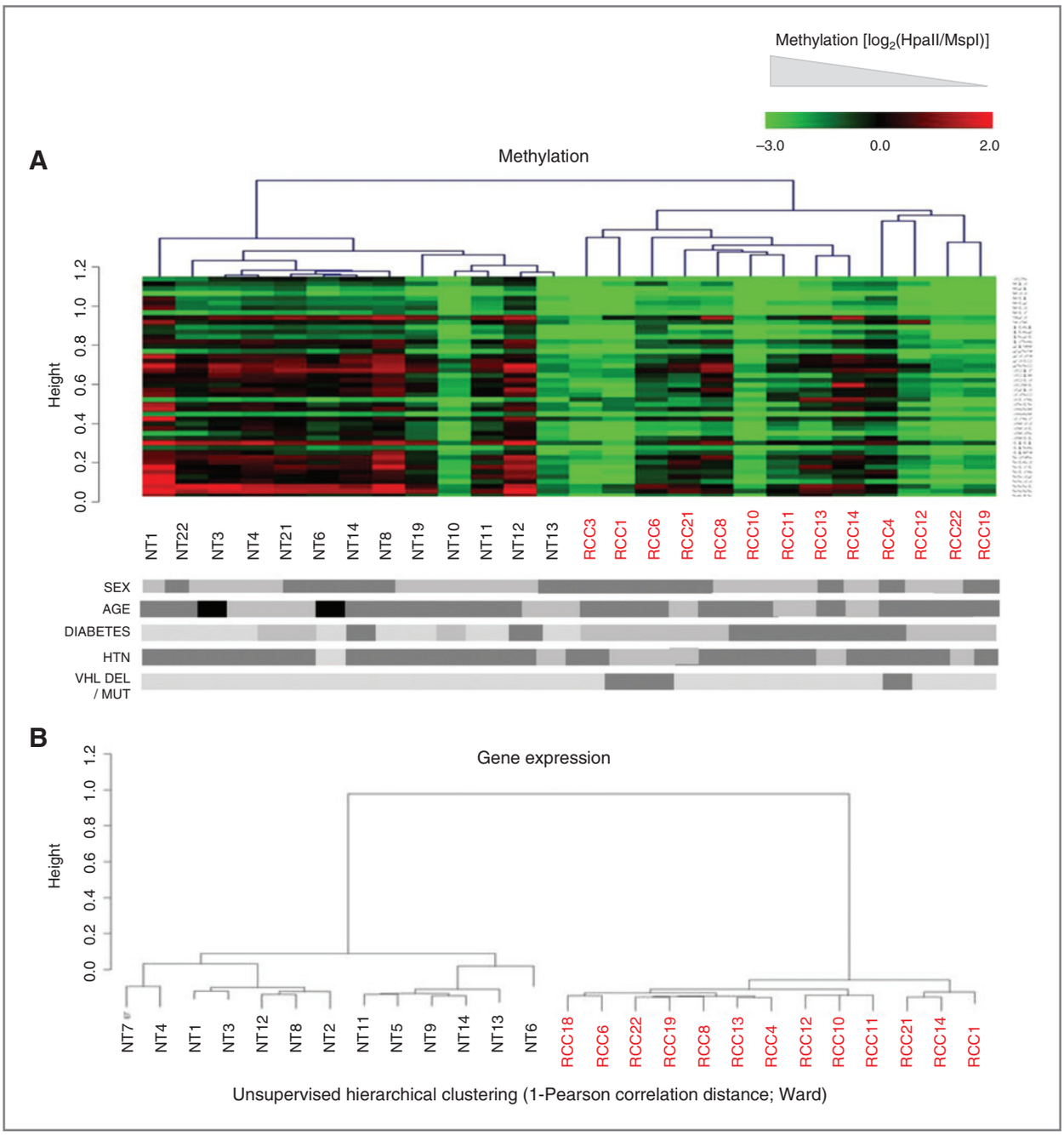

Figure 1. Methylation profiling separates RCC from normals. Methylation profiles generated by the HELP assay were used to cluster 13 RCC and 13 control samples by unsupervised hierarchical clustering (1-Pearson correlation distance and Ward agglomeration method). Heatmap shows hypermethylation (in green) and hypomethylation (in red) based on HpaII/MapI ratios. The controls formed a cluster that was distinct from RCC samples. No correlation with clinical characteristics was seen. Female gender, age 50- to 70-years-old, diabetes, hypertension, and VHL mutation were represented in gray. Male gender, age <50-years-old, no diabetes, no hypertension, no VHL deletion represented by light gray boxes. A, black boxes, age >70-years-old. B, unsupervised clustering based on gene-expression profiles reveals transcriptional differences between controls and RCC samples.

Differential methylation in RCC preferentially occurs at enhancer regions associated with H3K4Me1 marks

Cytosine methylation levels can regulate gene transcription by not only binding to promoter regions, but also by affecting transcription factor binding at enhancer regions in the genome. Because we found differential methylation at a notable proportion of intragenic sites, we analyzed their distribution within these regions and found differentially methylated regions (DMR) to preferentially involve introns (0.48% vs. 0.29%;

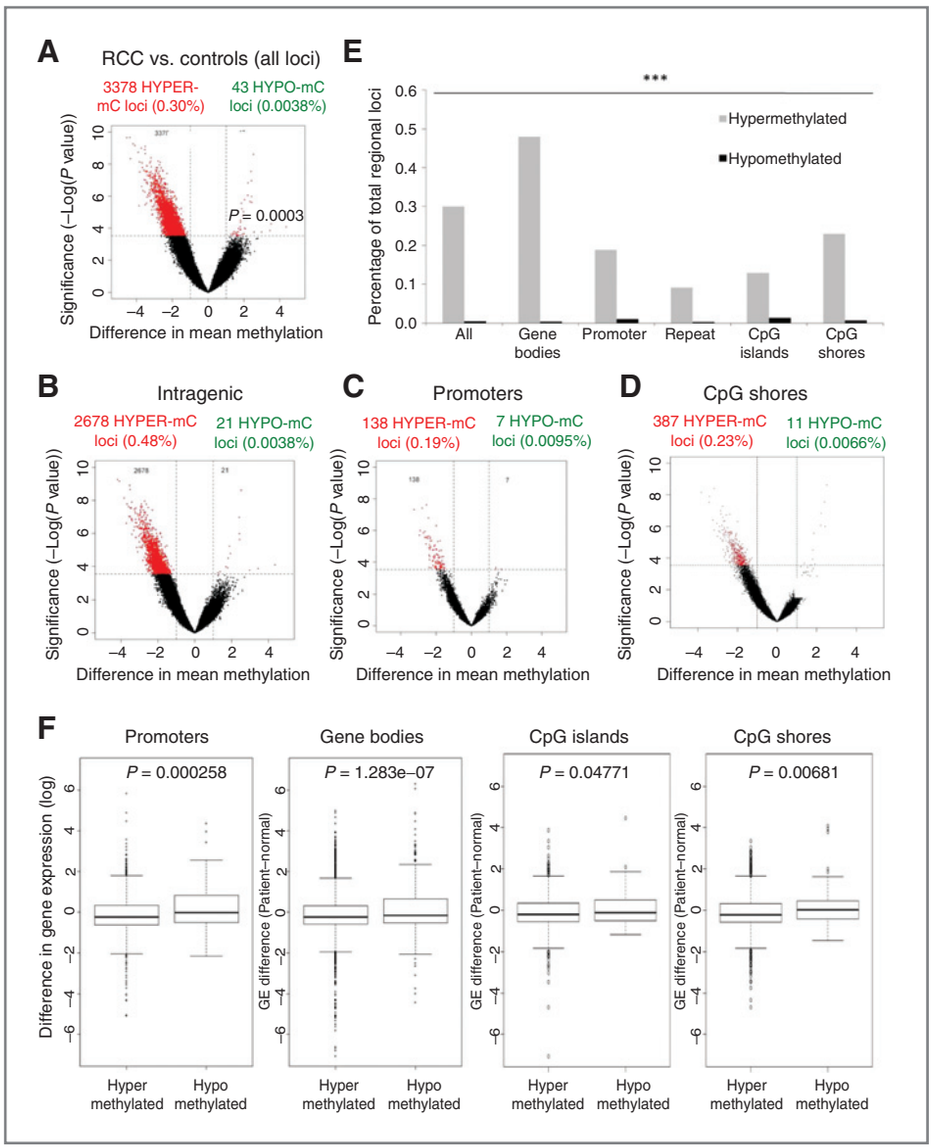

Figure 2. The majority of differentially methylated loci are hypermethylated in RCC and reside outside of CpG islands. A volcano plot is shown demonstrating the difference in mean methylation between all RCC samples and controls on the *x*-axis and the log of the *P* values between the means on the *y*-axis. A two-tailed *t* test was used to calculate the *P* values. A, significantly methylated loci (FDR < 0.1 by multiple testing) with a log fold change in mean methylation are labeled in red. B to D, volcano plots for gene body, promoter, and CpG shore loci also reveal mostly hypermethylated loci with highest percentages in gene bodies and CpG shores. E, the percentage of differentially methylated loci reveals significantly more hypermethylation genome wide (*t* test, $P < 0.001$). Difference of mean gene expression between RCC and controls is shown as box plots for hypermethylated and hypomethylated loci for all genomic locations. ***, $P < 0.001$; *t* test. F, hypermethylation is significantly associated with decrease in gene expression ($P < 0.01$).

$P < 0.00001$; Fig. 3A). Because intronic regions can contain enhancers, we wanted to evaluate the relationship of DMRs to other epigenetic marks that localize to regulatory regions of the genome. We compared DMRs with a panel of kidney-specific histone modifiers (H3K4me1, H3K4me3, and H3K36me3) obtained from chromatin immunoprecipitation followed by sequencing (CHIPSeq) analysis (31). We observed that DMRs associated with differentially expressed genes were preferentially localized to H3K4me1 marked regions of the genome (Fig. 3B) with a more than 3-fold enrichment (15.1% DMRs overlap vs. 4.2% for the array overlap with H3K4me1 marks from the BI27 dataset; $P < 0.05$; proportions test). H3K4me1 is a histone modification mark that can be observed in enhancer regions. Various important genes were aberrantly methylated at H3K4me1 marked regions and were underexpressed in RCC (Supplementary Table S2). Examples of multiple loci from the SMAD6 (Fig. 3C) and RXR-alpha (Supplementary Fig. S2) genes that were differentially hypermethylated in RCC overlap with H3K4me1 peaks generated from chromatin maps of adult kidney (BI27; ref. 31). Most of the differentially methylated loci were intronic and hypermethylation correlated with reduced gene expression of these genes in RCC. Importantly, comparison with chromatin occupancy maps from other tissue types (embryonic stem cells, hematopoietic cells, etc., shown at the bottom of the Fig. 3C) did not reveal increased numbers of non-kidney enhancer elements (marked by yellow) that coincided with differentially methylated loci in RCC.

Validation of these changes in an independent large dataset (TCGA) demonstrated that intronic loci for the SMAD6 gene were significantly methylated in RCC samples (Fig. 4A). Also, treatment of RCC-derived CAKI cells with

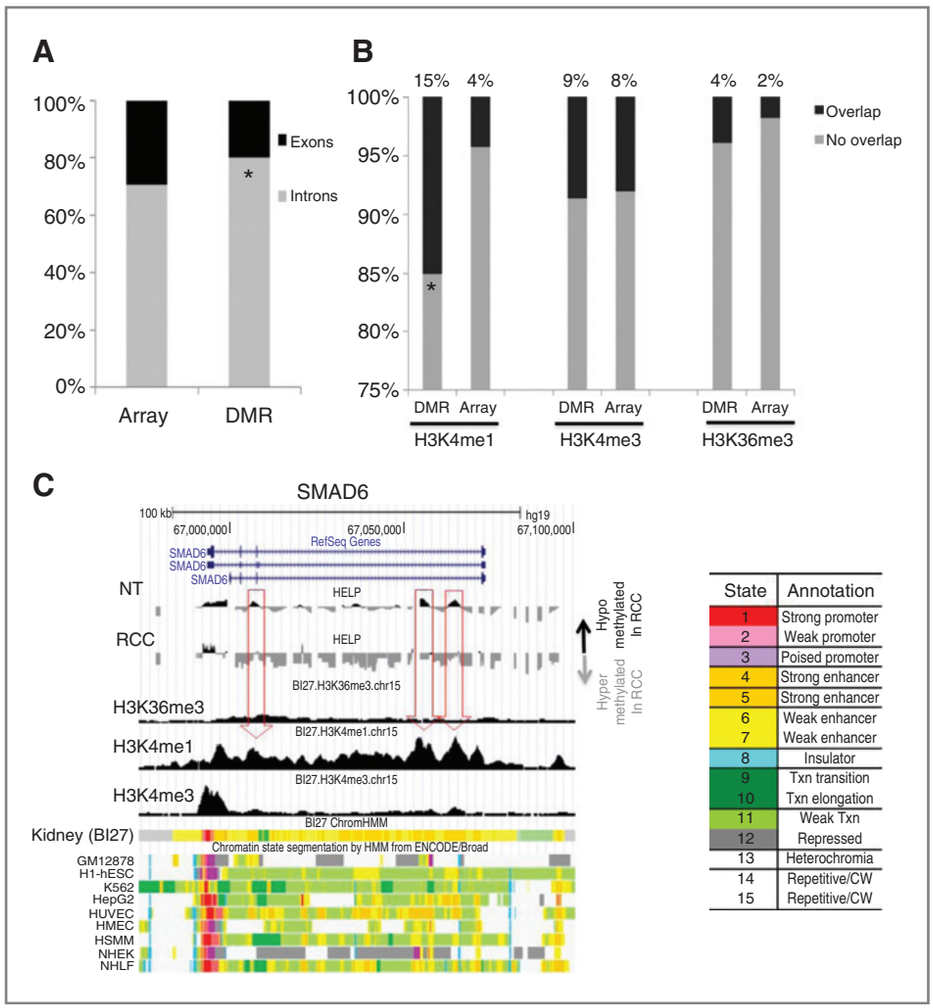

Figure 3. Differential methylation in RCC preferentially occurs at enhancer regions associated with H3K4Me1 marks. A, intragenic, DMRs were significantly located in intronic regions ($P < 0.05$ when compared with location of the whole array, test of proportions). B, overlap with adult kidney ChIP-seq data shows significant overlap of H3K4me1 marks with the DMRs ($P < 0.05$, test of proportions when compared with overlap with whole array). C, aberrantly methylated regions in the SMAD6 locus overlap with the kidney H3K4me1 peaks. Methylation values generated by the HELP assay [log (HpaII/MspI)] are shown as peaks (positive values corresponding with less methylation are shown in black and negative values corresponding to increased methylation are shown in gray). Histone modification peaks are shown for adult kidney (BI27). Chromatin occupancy for other tissue types is shown at bottom. Legend shows color coding for different regions with yellow regions representing enhancers. *, $P < 0.05$; test of proportions. GM12878, lymphoblastoid cell; H1-hESC, embryonic stem cells; K562, leukemic cells; HepG2, hepatic cells; HUVEC, umbilical vein endothelial cells; HMEC, mammary epithelial cells; HSMM, skeletal muscle myoblasts; NHEK, epidermal keratinocytes; NHLF, lung fibroblasts.

the DNMT inhibitor led to significant increase in SMAD6 expression, demonstrating that it is epigenetically regulated (Fig. 4B). Furthermore, SMAD6 was also significantly underexpressed in RCC samples when compared with controls when examined by RNA-seq in the TCGA dataset (Fig. 4C; *t* test; $P = 0.01$). Finally, the lower expression of SMAD6 was significantly associated with an adverse prognosis in multivariate analysis after adjusting for tumor stage (log-rank $P = 0.016$; Fig. 4D). In summary, these results indicate that the majority of DMRs in RCC are localized to gene body regions and a notable proportion are located on enhancer regions of the kidney genome, indicating that cytosine modification may act as a modifier of gene expression upon transcription factor binding.

Differential methylation in RCC displays specific genomic characteristics

Because we found that DMRs are enriched in gene regulatory elements, specifically enhancers, next, we wanted to determine whether the aberrantly methylated loci shared any common DNA elements. Because enhancers can bind transcription factors, we performed a search for transcription factor–binding sites enriched in these regions (CEAS program; ref. 26). Significant overrepresentation of binding sites for AP2α, AHR, HAIRY, and other transcription factors was seen in RCC (Supplementary Table S3). Interestingly, a number of transcription factors that are driven by hypoxia signaling (AHR, ARNT, and HIF1) were found to be enriched at these sites. The involvement of hypoxia signaling pathways reinforces the link to VHL–HIF pathway dysregulation that has been well described in RCC.

Next, we analyzed the gene pathways that were underexpressed and aberrantly hypermethylated in RCC and observed enrichment for cancer-associated genes (Supplementary Fig. S3 and Supplementary Table S4). Genes that were epigenetically silenced included novel candidates (SMAD7, SMAD6, HDAC9, and CSF1R) as well as those previously implicated (RB1 and TP53) in the pathogenesis of RCC.

We also analyzed the spatial distribution of hypermethylation in the genome and observed strikingly widespread involvement of all chromosomal regions (Fig. 6A)

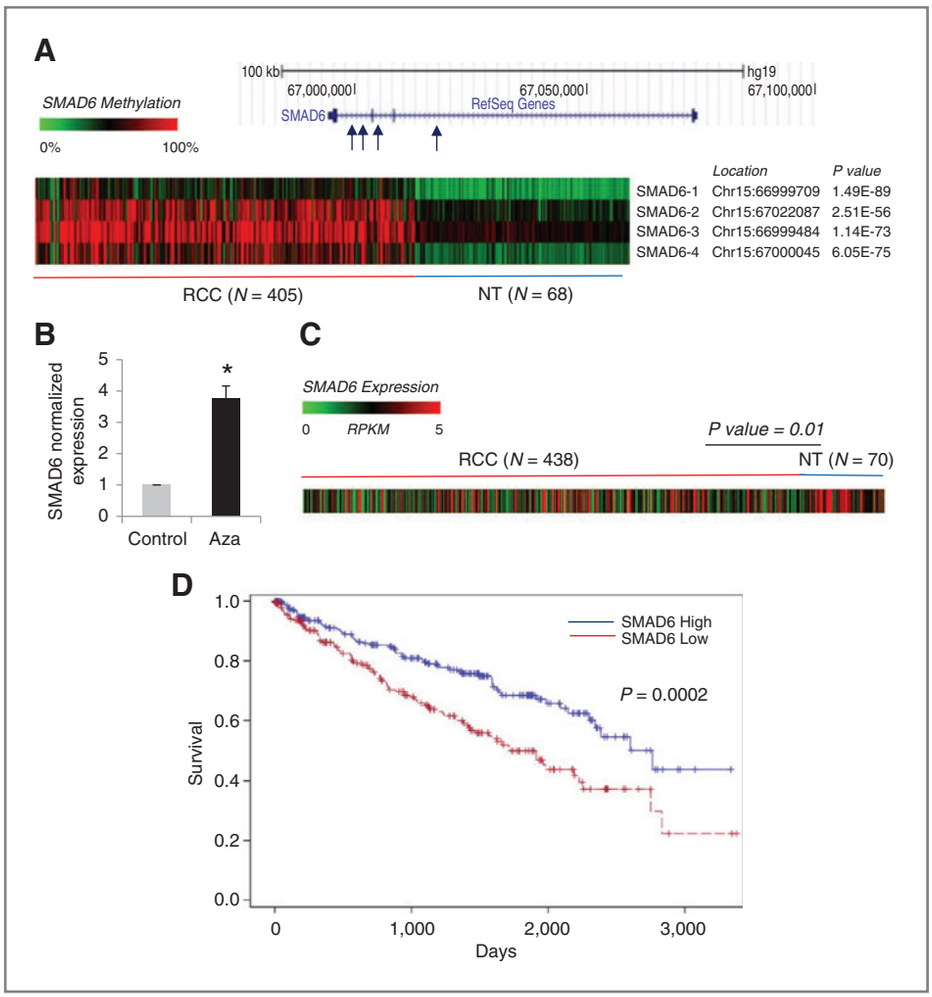

Figure 4. Aberrant methylation of intronic regions and reduced expression of SMAD6 are seen in an independent cohort of RCC samples and is predictive of adverse prognosis. A, methylation analysis of four intronic probes (arrows) from 405 RCC samples and controls reveals increased methylation in RCC samples. SMAD6 expression was assessed in CAKI cells after 5 days of exposure to 0.5 μmol/L 5-azacytidine. B, mean fold change over control is shown for two independent experiments, *t* test, $P = 0.02$. *, $P < 0.05$; *t* test. C, SMAD6 expression in this cohort is also significantly decreased in RCC samples when evaluated by RNA-seq. Kaplan–Meier survival curves of OS of 405 patients with RCC were plotted. D, blue solid lines, OS of patients with a higher expression of SMAD6 (top 20 percentile), whereas red dotted lines, OS of patients with a lower expression (lower 20 percentile; log-rank *P* value of 0.0002).

demonstrating the extent of aberrant hypermethylation in RCC.

Aberrant enhancer hypermethylation is prognostic in a large independent cohort of patients

To determine whether the aberrant hypermethylation seen in RCC has prognostic implications, we used the recently published TCGA dataset (20) comprising of 405 cases of RCC. Genomic locations of probes in the TCGA methylation dataset were matched to either promoters, gene bodies, and enhancers. A cumulative methylation score for each of these genomic locations was calculated for each patient when compared with normal reference obtained from 68 nontumor (NT) kidney controls. Correlation with OS (Kaplan–Meier analysis) demonstrated that increased amounts of methylation were predictive of adverse prognosis for all genomic locations in a univariate analysis (Fig. 5 and Supplementary Fig. S4; log-rank $P < 0.05$). Importantly, only methylation at enhancers was highly prognostic for survival in even a multivariate analysis that adjusted for tumor stage ($P = 0.02$), thus demonstrating the prognostic importance of enhancer associated epigenetic marks in RCC.

RCC cells are sensitive to low doses of the DNMT inhibitor

To further determine the functional relevance of the aberrant genome-wide hypermethylation observed in RCC, we tested the sensitivity of these cells to the DNMT inhibitor decitabine. RCC cell lines 786-O and 769-O and immortalized healthy renal tubular cells (HK-2 and HKP-8) were treated with low doses of decitabine. BrdUrd assay demonstrated that RCC cells were sensitive to growth inhibitory effects of a low dose of decitabine when compared with controls, thus demonstrating the therapeutic potential of targeting hypermethylation in RCC (Fig. 5C).

Widespread recurrent novel copy-number alterations can also be visualized by the HELP assay in RCC and are validated in a large independent cohort

Because MspI representations from the HELP array can yield copy-number data (22), we used these data to determine deletions and amplifications in RCC. Comparison of RCC and controls revealed a number of commonly deleted and amplified chromosomal regions in RCC. These included novel areas of deletion (chr13q12, 16q21, 14q13, and

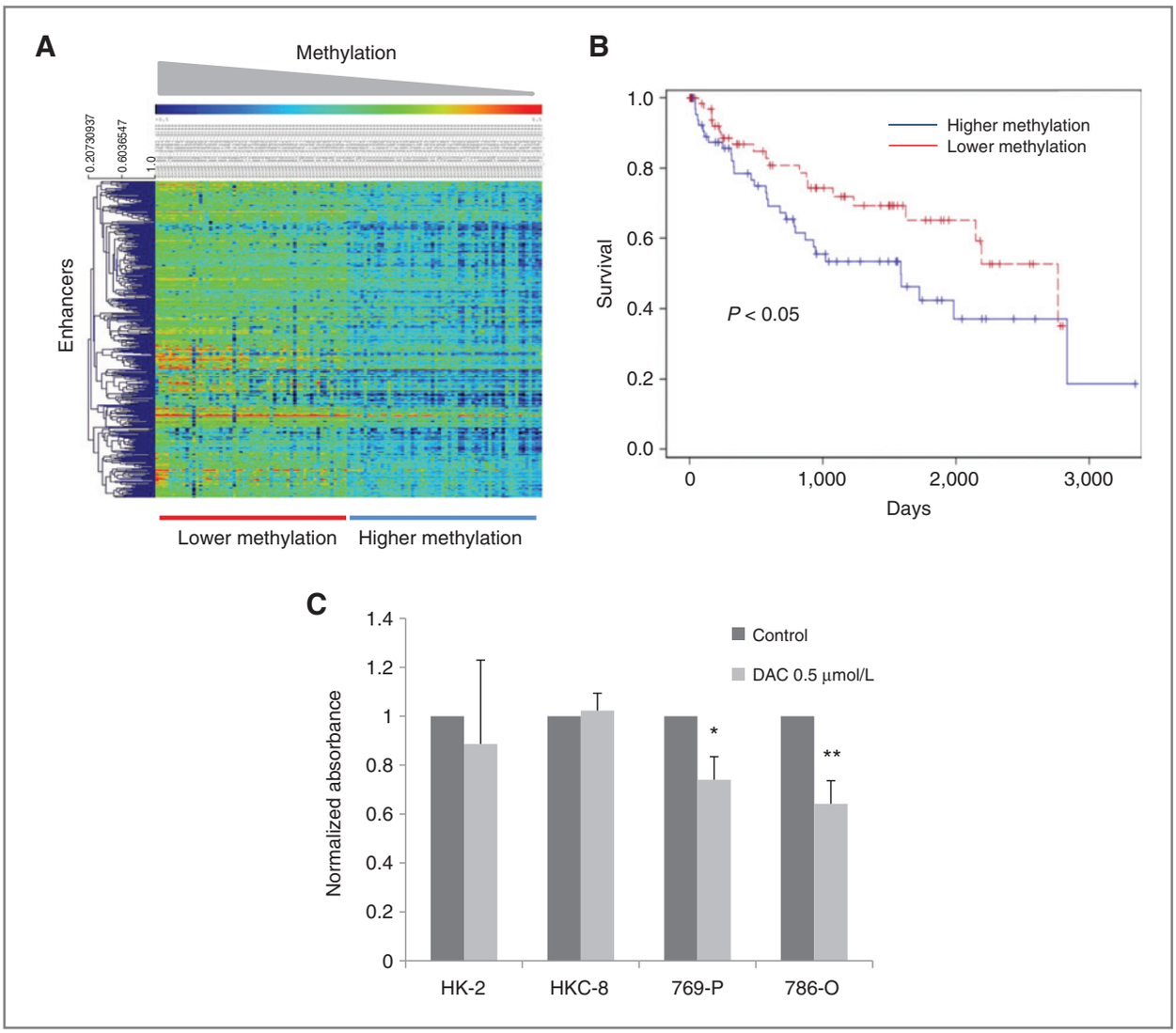

Figure 5. Aberrant enhancer hypermethylation is prognostic in the large independent cohort of patients, and RCC is sensitive to low doses of DNMT inhibitors. Heatmap of the respective patients (horizontal order) and the top 500 enhancer loci (vertical order). Patients are ranked in descending order based on the methylation. A, Kaplan–Meier survival curves of OS of 405 patients with RCC were plotted. B, blue solid lines, OS of patients with a higher methylation (top 20 percentile) of kidney-specific enhancers, whereas red dotted lines represent OS of patients with a lower methylation (lower 20 percentile; log-rank *P* value of 0.03 for univariate and 0.03 for multivariate analysis). C, RCC cell lines (769-P and 786-O) and healthy kidney cells (HK-2 and HKC-8) were grown for 5 days in the presence and absence of 0.5 μmol/L decitabine, and BrdUrd incorporation was measured and normalized. Significant growth inhibition was seen in RCC cells (two-tailed *t* test, *P* value of < 0.05). *, $P < 0.05$; **, $P < 0.01$, *t* test.

others) and amplification (chr 17p13, 12p13, 16q24, and others) that have not been seen by previous studies (Fig. 6A). Interestingly, the chromosome 3p24-25 region that includes the VHL gene was found to be deleted in three samples (Supplementary Fig. S5; ref. 32), thus confirming the applicability of our findings to other patient cohorts. Several important genes not previously implicated in RCC were found to be affected by deletions [large tumor suppressor (LATS), CDC25A, SMARCC1, and others; Supplementary Table S5] and amplifications [PDGFα (platelet-derived growth factor α), SOX18, and others; Supplementary Table S6], thus providing a high-resolution genomic map of CNVs in RCC. We validated some of these novel copy-number changes by genomic qPCR and observed that the deletion of LATS (Fig. 6B and E) and amplification of PDGFα (Fig. 6C and D) were observed and correlated strongly with the findings of the array. These changes were also validated in the independent TCGA dataset. LATS1 was found to be deleted in 121 of 405 (30%) cases and PDGFα was found to be amplified in 151 of 405 (37%) cases, providing independent validation of our findings. Furthermore, the expression of PDGFα was significantly increased in RCC samples when compared with normal tumor kidney controls ($P = 0.01$; *t* test),

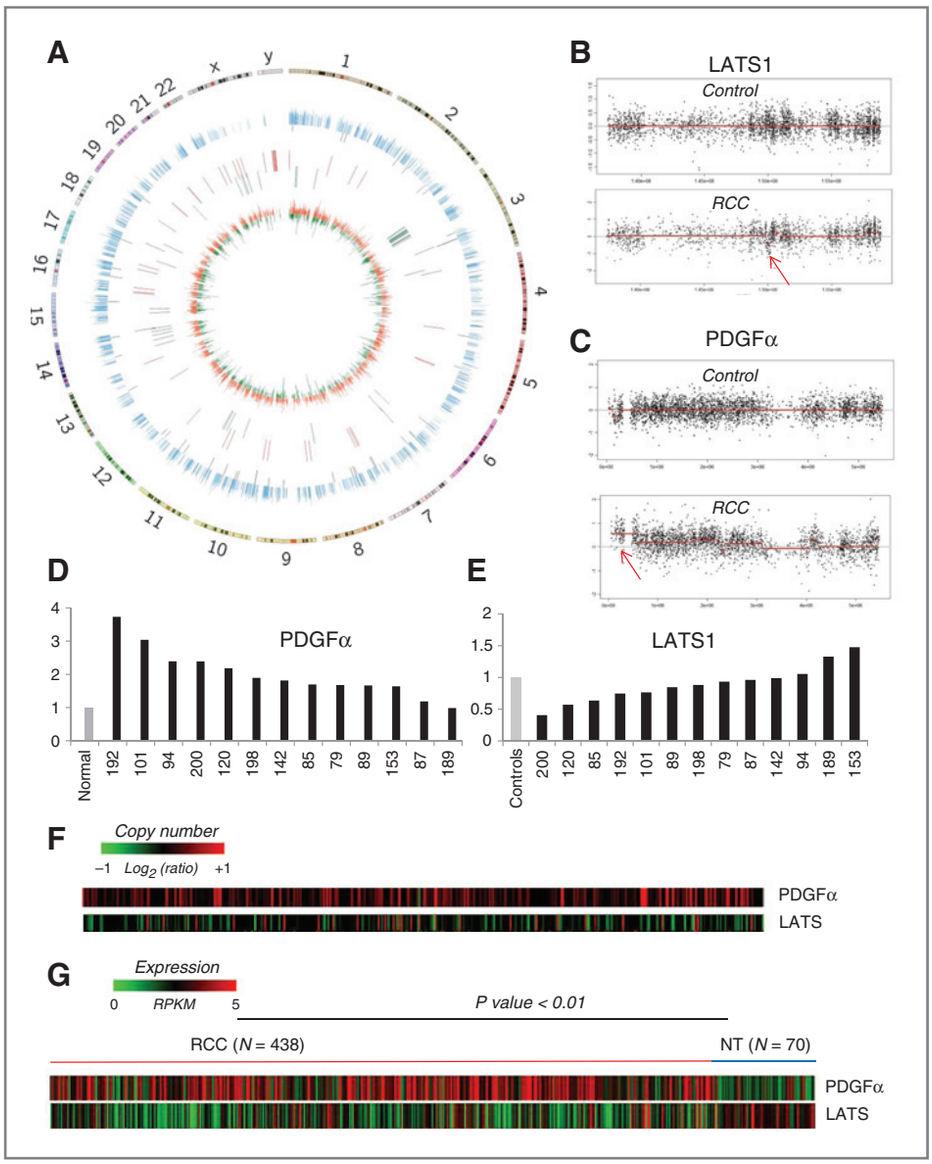

Figure 6. MspI representations can detect CNVs in RCC that are validated in an independent cohort. CNV information can be detected through the MspI representation in the HELP assay. A, a circos plot with methylation (hypermethylation in blue, outer ring), CNV (middle ring), and expression (underexpressed in green, innermost ring) is shown. Deletions represented in green and amplifications in red are shown for all 13 patients with RCC. An example of deletion of LATS (B) and amplification of PDGFα (C) is shown for control and RCC samples using MspI representations. Validation by qPCR shows amplification of PDGFα (D) and deletion of the LATS gene (E) is seen in RCC when compared with pooled normal controls. *X*-axis represents normalized fold change. F, PDGFα is found to be amplified in 151 of 405 cases and LATS is found to be deleted in 121 of 405 cases in the TCGA dataset. G, RNA-seq expression data reveal overexpression of PDGFα and reduced expression of LATS in TCGA RCC samples when compared with nontumor controls ($n = 68$) as shown by heatmaps.

whereas LATS expression was found to be significantly decreased in RCC (*t* test; $P = 0.01$; Fig. 6G).

Discussion

Advanced RCC is a malignancy that has a poor prognosis and limited treatment options. It is well known that there are a number of genetic abnormalities that result in the transformation of normal tissue to RCC. More recently, epigenetic methylation changes at promoters have been shown to turn off certain tumor-suppressor genes in RCC, suggesting a role of aberrant methylation in the pathobiology of this disease (33, 34). We used an unbiased, high resolution, global assay to look for epigenomic disturbances in RCC. Our studies revealed widespread aberrations in DNA methylation that clearly distinguished RCC from normal controls. Furthermore, a large proportion of aberrant methylation affected kidney-specific enhancers. These findings demonstrate that gene body methylation seen by us and others can affect regulatory regions of the genome and has prognostic implications.

Our epigenetic studies were based on a high-resolution version of the HELP assay that examines cytosine methylation at 1.3 million CCGG (HpaII) sites and is not biased to areas of high CpG density (14). This assay allowed us to examine RCC for changes outside of promoters and CpG islands, areas that have been the focus of most cancer epigenetic studies previously. In fact, we found that the majority of common differentially hypermethylated cytosines in RCC samples were not located in promoters and CpG islands. Recent work has also shown that non-CpG island cytosine methylation can be

important in controlling gene transcription and can be involved in normal development and carcinogenesis (18, 25). We show that gene bodies and CpG shores are regions characterized by a higher proportion of aberrant hypermethylation in RCC.

A notable number of differentially methylated loci were found to occur in enhancer regions of the genome. A comparison of DMRs with genome-wide maps of H3K4me1 localizations from adult kidney tissues revealed that hypermethylation occurred at these intronic regions as illustrated for SMAD6 and other genes. Recent studies have started to show that DNA methylation can occur at regulatory regions of the genome and affects transcription factor binding, thereby regulating gene transcription. Our data show this phenomenon for the first time in renal cancer. Most importantly, we demonstrated that increased methylation of enhancers in an independent large TCGA dataset was a marker of adverse prognosis. The TCGA report had shown that increased methylation was seen in RCC and had correlated it with disease stage. We demonstrated that enhancer methylation was predictive of decreased survival even after adjusting for disease stage and was, therefore, an independent risk factor in RCC. Other recent studies have also correlated the presence of aberrant histone methylation with worse prognosis and support the role of epigenetic alterations in disease progression and response to treatment (35–37).

Furthermore, as enhancer regions are hotspots of transcription factor binding, we conducted an unbiased search of motif analysis of DMRs and demonstrated that the DMRs were enriched for binding sites for transcription factors involved in hypoxia pathways. Hypoxia pathways have been well studied in RCC and mutations in the VHL gene have been well described in this malignancy (38). Our findings demonstrate that transcription factors regulated by alterations in the VHL–HIF pathway may play a role in driving epigenetic changes that are seen in RCC and in turn may influence the expression of oncogenic pathways through this epigenetic modification. We also observed that aberrant methylation of intronic enhancers of the SMAD6 gene was associated with decreased expression of this gene in RCC. SMAD6 is an inhibitor of the TGFβ receptor I kinase and reduced expression of SMAD6 can lead to overactivation of the TGFβ pathway. Increased TGFβ signaling has been seen in renal cell cancer and is associated with worse prognosis (39). We also observed that lower expression of SMAD6 is associated with worse prognosis and SMAD6 underexpression provides a potential mechanism for increased TGFβ signaling seen in RCC.

Hypomethylating agents such as decitabine and 5-azacytidine are inhibitors of DNMTs and currently used in the treatment of myelodysplastic syndrome (40). These agents are rarely used in solid tumors, as most of these tumors are characterized by global hypomethylation (41). However, our study demonstrates significant global hypermethylation in RCC tumor samples and reveals that alterations in DNA methylation are also demonstrated, which raises the possibility that these tumors can be targeted by hypomethylating agents (42). The sensitivity of both VHL-mutant and -nonmutant RCC cell lines to the growth inhibitory effects of decitabine reinforces the therapeutic potential of these agents in RCC. Hypermethylation has been seen in other tumors such as acute myeloid leukemia and myelodysplastic syndromes and some of these are associated with inactivating mutations in demethylating ten-eleven translocation (TET) enzymes. Even though mutations in TET and isocitrate dehydrogenase enzymes are not seen in RCC, it is possible that these tumors have transcriptional dysregulation of members of the demethylation pathway that needs to be evaluated in future studies. In summary, we present a high-resolution epigenomic and genomic map of RCC and demonstrate that these high-throughput assays have the potential of increasing our understanding of renal cell tumorigenesis. The data generated under this project will be made publicly available for the wider community to further analyze and understand RCC development.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: C.Y. Hu, J. Pullman, J.M. Greally, K. Susztak, A. Verma

Development of methodology: M.C. Izquierdo, T.D. Bhagat, M. Suzuki, J. Pullman, A. Verma

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D. Mohtat, N. Shenoy, S. Bhattacharya, A.S.D. Park, O. Giricz, N. Vallumsetla, K. Gundabolu, K. Ware, J. Pullman, K. Susztak, A. Verma

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): C.Y. Hu, D. Mohtat, Y. Yu, Y.-A. Ko, N. Shenoy, A.S.D. Park, X.S. Liu, A. Verma

Writing, review, and/or revision of the manuscript: C.Y. Hu, T.D. Bhagat, J.M. Greally, K. Susztak, A. Verma

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C.Y. Hu, O. Giricz, X. S. Liu, K. Susztak, A. Verma,

Study supervision: T.D. Bhagat, A. Verma

Grant Support

This work was supported by the NIH (R01DK087635), the Hershaft Family Foundation, the Leukemia and Lymphoma Society, and the Department of Defense.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 27, 2014; revised May 12, 2014; accepted May 15, 2014; published OnlineFirst June 10, 2014.

References

1. Alvarez H, Opalinska J, Zhou L, Sohal D, Fazzari MJ, Yu Y, et al. Widespread hypomethylation occurs early and synergizes with gene amplification during esophageal carcinogenesis. PLoS Genet 2011;7:e1001356.

2. De S, Shaknovich R, Riester M, Elemento O, Geng H, Kormaksson M, et al. Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. PLoS Genet 2013;9:e1003137.
3. Suzuki M, Oda M, Ramos MP, Pascual M, Lau K, Stasiek E, et al. Late-replicating heterochromatin is characterized by decreased cytosine methylation in the human genome. Genome Res 2011;21: 1833–40.
4. Steidl U, Steidl C, Ebralidze A, Chapuy B, Han HJ, Will B, et al. A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. J Clin Invest 2007;117:2611–20.
5. Lamprecht B, Walter K, Kreher S, Kumar R, Hummel M, Lenze D, et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. Nat Med 2010;16: 571–9.
6. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 2013;153:320–34.
7. Baldewijns MM, van Vlodrop IJ, Schouten LJ, Soetekouw PM, de Bruine AP, van Engeland M. Genetics and epigenetics of renal cell cancer. Biochim Biophys Acta 2008;1785:133–55.
8. Cohen HT, McGovern FJ. Renal-cell carcinoma. N Engl J Med 2005;353:2477–90.
9. Gordan JD, Lal P, Dondeti VR, Letrero R, Parekh KN, Oquendo CE, et al. HIF-alpha effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma. Cancer Cell 2008;14:435–46.
10. Powles T, Chowdhury S, Jones R, Mantle M, Nathan P, Bex A, et al. Sunitinib and other targeted therapies for renal cell carcinoma. Br J Cancer 2011;104:741–5.
11. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature 2011;469:539–42.
12. Murtaugh LC, Stanger BZ, Kwan KM, Melton DA. Notch signaling controls multiple steps of pancreatic differentiation. Proc Natl Acad Sci U S A 2003;100:14920–5.
13. Esteller M. Cancer epigenetics: DNA methylation and chromatin alterations in human cancer. Adv Exp Med Biol 2003;532:39–49.
14. Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. Cancer Cell 2010;17: 13–27.
15. Khulan B, Thompson RF, Ye K, Fazzari MJ, Suzuki M, Stasiek E, et al. Comparative isoschizomer profiling of cytosine methylation: the HELP assay. Genome Res 2006;16:1046–55.
16. Thompson RF, Reimers M, Khulan B, Gissot M, Richmond TA, Chen Q, et al. An analytical pipeline for genomic representations used for cytosine methylation studies. Bioinformatics 2008;24:1161–7.
17. Arai E, Kanai Y. Genetic and epigenetic alterations during renal carcinogenesis. Int J Clin Exp Pathol 2010;4:58–73.
18. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009;41:178–86.
19. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009;27:361–8.
20. Creighton CJ, Morgan M, Gunaratne PH, Wheeler DA, Gibbs RA, Robertson A, et al.Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013;499:43–9.
21. Woroniecka KI, Park AS, Mohtat D, Thomas DB, Pullman JM, Susztak K. Transcriptome analysis of human diabetic kidney disease. Diabetes 2011;60:2354–69.
22. Oda M, Glass JL, Thompson RF, Mo Y, Olivier EN, Figueroa ME, et al. High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. Nucleic Acids Res 2009;37:3829–39.
23. Anguiano A, Tuchman SA, Acharya C, Salter K, Gasparetto C, Zhan F, et al. Gene expression profiles of tumor biology provide a novel approach to prognosis and may guide the selection of therapeutic targets in multiple myeloma. J Clin Oncol 2009;27:4197–203.
24. Figueroa ME, Reimers M, Thompson RF, Ye K, Li Y, Selzer RR, et al. An integrative genomic and epigenomic approach for the study of transcriptional regulation. PLoS ONE 2008;3:e1882.
25. Figueroa ME, Wouters BJ, Skrabanek L, Glass J, Li Y, Erpelinck-Verschueren CA, et al. Genome-wide epigenetic analysis delineates a biologically distinct immature acute leukemia with myeloid/T-lymphoid features. Blood 2009;113:2795–804.
26. Shin H, Liu T, Manrai AK, Liu XS. CEAS: cis-regulatory element annotation system. Bioinformatics 2009;25:2605–6.
27. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9.
28. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 2011;473:43–9.
29. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods 2012;9:215–6.
30. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet 2002;3:415–28.
31. Aiden AP, Rivera MN, Rheinbay E, Ku M, Coffman EJ, Truong TT, et al. Wilms tumor chromatin profiles highlight stem cell properties and a renal developmental network. Cell Stem Cell 2010;6:591–602.
32. Cowey CL, Rathmell WK. VHL gene mutations in renal cell carcinoma: role as a biomarker of disease outcome and drug efficacy. Curr Oncol Rep 2009;11:94–101.
33. Morris MR, Ricketts CJ, Gentle D, McRonald F, Carli N, Khalili H, et al. Genome-wide methylation analysis identifies epigenetically inactivated candidate tumour-suppressor genes in renal cell carcinoma. Oncogene 2011;30:1390–401.
34. Morris MR, Ricketts C, Gentle D, Abdulrahman M, Clarke N, Brown M, et al. Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. Oncogene 2010;29:2104–17.
35. Mosashvilli D, Kahl P, Mertens C, Holzapfel S, Rogenhofer S, Hauser S, et al. Global histone acetylation levels: prognostic relevance in patients with renal cell carcinoma. Cancer Sci 2010;101:2664–9.
36. Minardi D, Lucarini G, Filosa A, Milanese G, Zizzi A, Di Primio R, et al. Prognostic role of global DNA-methylation and histone acetylation in pT1a clear cell renal carcinoma in partial nephrectomy specimens. J Cell Mol Med 2009;13:2115–21.
37. Kanao K, Mikami S, Mizuno R, Shinojima T, Murai M, Oya M. Decreased acetylation of histone H3 in renal cell carcinoma: a potential target of histone deacetylase inhibitors. J Urol 2008;180:1131–6.
38. Pantuck AJ, Zeng G, Belldegrun AS, Figlin RA. Pathobiology, prognosis, and targeted therapy for renal cell carcinoma: exploiting the hypoxia-induced pathway. Clin Cancer Res 2003;9:4641–52.
39. Casalena G, Daehn I, Bottinger E. Transforming growth factor-beta, bioenergetics, and mitochondria in renal disease. Semin Nephrol 2012;32:295–303.
40. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, Santini V, Finelli C, Giagounidis A, et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. Lancet Oncol 2009;10:223–32.
41. Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene 2002;21:5400–13.
42. Negrotto S, Hu Z, Alcazar O, Ng KP, Triozzi P, Lindner D, et al. Noncytotoxic differentiation treatment of renal cell cancer. Cancer Res 2011;71:1431–41.