Altered DNA Methylation Landscapes of Polycomb-Repressed Loci Are Associated with Prostate Cancer Progression and ERG Oncogene Expression in Prostate Cancer

Ken Kron\textsuperscript{1,2}, Dominique Trudel\textsuperscript{3}, Vaijayanti Pethe\textsuperscript{1}, Laurent Briollais\textsuperscript{1}, Neil Fleshner\textsuperscript{4}, Theodorus van der Kwast\textsuperscript{2,3}, and Bharati Bapat\textsuperscript{1,2}

Abstract

\textbf{Purpose:} To assess differentially methylated “landscapes” according to prostate cancer Gleason score (GS) and ERG oncogene expression status, and to determine the extent of polycomb group (PcG) target gene involvement, we sought to assess the genome-wide DNA methylation profile of prostate cancer according to Gleason score and ERG expression.

\textbf{Experimental Design:} Genomic DNA from 39 prostate cancer specimens was hybridized to CpG island microarrays through differential methylation hybridization. We compared methylation profiles between Gleason score and ERG expression status as well as Gleason score stratified by ERG expression status. In addition, we compared results from our dataset to publicly available datasets of histone modifications in benign prostate cells.

\textbf{Results:} We discovered hundreds of distinct differentially methylated regions (DMR) associated with increasing Gleason score and ERG. Furthermore, the number of DMRs associated with Gleason score was greatly expanded by stratifying samples into ERG-positive versus ERG-negative, with ERG-positive/GS-associated DMRs being primarily hypermethylated as opposed to hypomethylated. Finally, we found that there was a significant overlap between either Gleason score–related or ERG-hypermethylated DMRs and distinct regions in benign epithelial cells that have PcG signatures (H3K27me3, SUZ12) and lack active gene expression signatures (H3K4me3, RNA pol II).

\textbf{Conclusions:} This work defines methylation landscapes of prostate cancer according to Gleason score, and suggests that initiating genetic events may influence the prostate cancer epigenome, which is further perturbed as prostate cancer progresses. Moreover, CpG islands with silent chromatin signatures in benign cells are particularly susceptible to prostate cancer–related hypermethylation.

Clinical Cancer Research; 19(13); 3450–61. ©2013 AACR.

Introduction

Prostate cancer remains a significant health burden for men and represents a challenge for physicians and patients in deciding optimal treatment course as overdiagnosis has become more prevalent (1). Much of the challenge can be attributed to the large number of indolent tumors identified in recent years due to widespread prostate-specific antigen (PSA) testing (2, 3). To address this challenge, a great deal of effort is focused on searching the genome for biomarkers that may better inform of prostate cancer prognosis (4). Despite this effort, there has been a general deficiency of clinically useful biomarkers that can be applied pre- and posttreatment, as most do not offer significant additional information beyond classic pathologic parameters such as grade [as indicated by Gleason score (GS)] and stage. To reflect the morphologic heterogeneity of a prostate cancer, the Gleason score (ranging from 2 to 10) is composed of the two most dominant architectural patterns, which are numbered 1 to 5. If only one pattern is present, this is doubled to give the Gleason score (5, 6).

Previous studies have attempted to discern genetic and expression-based markers that may be indicators of prostate cancer prognosis, either correlating with clinical variables such as Gleason score or predicting disease recurrence (7, 8).
Numerous genetic and epigenetic changes are associated with prostate carcinogenesis and disease progression, including expression of an oncogenic fusion gene (TMPRSS2:ERG) as an early genetic event and increased DNA methylation. Here, we find that stratifying prostate cancer by ERG expression status expands the potential number of prognosis-related epigenetic events. Furthermore, we find that many regions with increased methylation may already be in a transcriptionally silent state in benign prostate cells through epigenetic mechanisms other than DNA methylation. This work further characterizes molecular subtypes of prostate cancer and identifies potential new biomarkers associated with disease course within these subtypes.

After the initial discovery of TMPRSS2:ERG gene fusions as a frequent event in prostate cancer, several earlier studies suggested it was a marker of poor prognosis (9). Recent work, however, failed to confirm that TMPRSS2:ERG correlates with adverse tumor characteristics and unfavorable prognosis (10, 11). Nonetheless, this fusion is a likely initiating event in prostate cancer development representing a separate etiology with unique expression and epigenetic profiles (12–14).

Many earlier studies using primary prostate tissue have limited their analysis to a set of previously identified genes (15), whereas others have used in vitro epigenetic unmasking of methylated genes in cell lines followed by expression profiling (16). More recent work, however, has taken advantage of technologic advances in methylation profiling that allow increasing genomic coverage to identify prognostic methylation changes in tissue samples (17, 18). Previous DNA methylation profiling in prostate cell lines has also identified an increase in DNA methylation in PC3 cells versus benign prostate epithelial cells of genes repressed by histone 3, lysine 27 trimethylation (H3K27me3) via the polycomb-repressive complex 2 (PRC2; ref. 19). Others have identified DNA methylation changes that correspond to polycomb group (PcG) target genes in renal cancer (20) and early-stage liver cancer (21) but do not directly compare with polycomb-repressive marks in benign tissue of the same origin.

In this study, we sought to analyze methylation on a genome-wide scale to discover methylated loci associated with Gleason score in two manners—irrespective of ERG expression status or stratified by ERG expression—using an 18-gauge needle. DNA was extracted from this tissue using the QIAMP DNA Mini Kit (Qiagen) according to the manufacturer’s recommended protocol for tissue specimens. Nine patients with GS7 had both a clearly delineated Gleason pattern 3 and Gleason pattern 4 within the same tumor, which were dissected and extracted separately. For the remaining 10 GS7 specimens, only 1 dominant Gleason pattern was present in the frozen tumor, so DNA from the entire carcinoma was extracted.

We used the differential methylation hybridization (DMH) approach to enrich for methylated DNA regions as previously described (24). The pattern 3 and 4 specimens from 9 patients with GS7 were arrayed separately, with subsequent array results averaged to give a single reading for each patient. The GS6 and GS8 specimens were previously analyzed and have been deposited into Gene Expression Omnibus (GEO) as dataset GSE 15298.

CpG island microarray data analysis
Normalized log10 ratios of tumor/reference intensities from Agilent Feature Extraction software were subjected to a batch effect correction model as previously described (25). The corrected data was subsequently analyzed using SPSS v18. For the Gleason score analysis, we conducted ANOVA across 3 groups of GS6, GS7, and GS ≥ 8 and filtered
significant probes ($P < 0.05$) to have a difference of 2-fold or more when comparing the GS6 to GS $\geq$ 8 groups. When comparing GS6 versus GS7 and GS7 versus GS $\geq$ 8, fold change criteria was relaxed to include differences 1.5 or more, as fold changes 2 or more limited these datasets to 2 and 9 significant probes, respectively. For ERG-stratified methylation analysis, classical $t$ tests were conducted and significant probes ($P < 0.05$) were again filtered to have a difference 2-fold or more when comparing the 2 groups.

To assess differences between the locations of differentially methylated regions (DMR), we first assigned location using the annotate feature within CisGenome software v2.0 (The Johns Hopkins University, Baltimore, MD; ref. 26). Bed files of genome coordinates (hg18) were loaded into CisGenome and coordinates were then assigned to the closest gene and gene region based on the following criteria: "Promoter," $+2,500$ to $-1,500$ relative to transcription start site (TSS); "Intragenic," $-1,501$ TSS to transcription end site (TES); "Intergenic," $> +2,500$ TSS and beyond TES. Proportional differences in DMR location were then tested using the $\chi^2$ test and $P \leq 0.017$ (Bonferroni adjustment of 3 tests) were considered significant.

**ERG immunohistochemistry**

FFPE tumor samples were obtained and immunohistochemistry was conducted on the block matching the frozen tumor in 12 of 35 cases, whereas in the remaining cases tumor from an adjacent tissue block, but representative of the same tumor, was used. Immunostaining of the tissue sections for ERG was conducted as previously described (23).

A methodologic flow-chart of case selection and samples analyzed for DNA methylation and ERG immunohistochemistry is shown in Supplementary Fig. S1.

**Unsupervised clustering**

Clustering of data was conducted using normalized log ratios from CpG island microarrays with Cluster 3.0 software (University of Tokyo, Human Genome Center, Tokyo, Japan) and visualized using TreeView software v1.60 (Eisen Laboratory University of California at Berkeley, CA). As this software could not analyze all 237,000 probes, datasets were filtered for probes that had a SD across samples of 1.5 or more. This limited the Agilent CpG island methylation data to 193 probes. Unsupervised clustering was then conducted using the Euclidean distance similarity metrics for probes and cases and average linkage clustering method.

**Functional annotation**

The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used for functional annotation clustering of DMRs associated with specific genes. Gene lists were uploaded into the DAVID list manager using the official gene symbol and a background gene list based on the Agilent human CpG island platform was selected. Significant annotation clusters were based on an enrichment score 1.3 or more as described by Huang da and colleagues (27).

Publicly available ChIP-chip and methylation data analysis

Processed chromatin immunoprecipitation (ChIP)-chip datasets published by Takeshima and colleagues (28) and Gal-Yam and colleagues (18) were downloaded from the GEO website under accessions GSE15154 and GSE12334, respectively. The Takeshima dataset allowed direct probe comparisons with our methylation data as the Agilent CpG island platform was similarly used for ChIP-chip. For each histone modification (H3Ac, H3K4me3, H3K9me3, and H3K27me3) and RNA polymerase II (RNA pol II) binding analyzed in RWPE-1 cells as well as DNA methylation of other cell lines in the Takeshima dataset, we conducted $t$ test statistics comparing means from our regions of interest versus the average from all probes throughout the genome for each individual ChIP/methylation profile.

For validation of findings in our dataset, we used a set of 57 DNA methylation markers identified as associated with disease recurrence (biochemical, clinical, or systemic) from a recent publication using Illumina27 arrays (18). Agilent CpG island array probes or Nimblegen custom tiling probes from the ChIP-chip data that overlapped the CpGs analyzed by Illumina arrays for each of the 57 genes (18) were used for analysis of histone modifications and RNA pol II binding in benign cells. A Bonferroni corrected $t$ test was again used to analyze this data, with $P \leq 1.14 \times 10^{-3}$ (44 tests conducted) considered significant.

**Results**

**Gleason score–related methylation changes**

To discover genes that had significantly different methylation profiles across all Gleason scores (6, 7, and 8), we established criteria for differential methylation as a fold change between GS6 and GS $\geq 8$ of $\geq 2$ (positive or negative) and ANOVA $P$ value comparing all Gleason score of 0.05 or less. This produced a list of 474 differentially methylated probes, representing 427 unique gene regions across all Gleason scores analyzed (Figure 1B and Supplementary Table S1). All significant probes from this analysis were similarly significant in our previous analysis of GS6 versus GS8, which was conducted using a Bayesian model (24). Thus, we considered this approach to be slightly more stringent than our previous approach.

Interestingly, 321 (75.2%) of the DMRs were hypomethylated according to increasing Gleason score compared with 106 (24.8%) that were hypermethylated. We also conducted pairwise analyses comparing GS6 with GS7 and GS7 with GS $\geq 8$. Fold changes within these comparisons were considerably less than comparing GS6 with GS8, so we relaxed the significance criteria to include fold changes 1.5 or more. In the GS6 versus GS7 analysis, 189 probes representing 182 distinct genomic locations were differentially methylated (Supplementary Table S1). The majority of these regions were hypomethylated in GS7 compared with GS6 (167), compared with just 15 that were hypermethylated. The most significant hypomethylated region was found in an intergenic location adjacent to BARHL2.
Fold change compared to control sample as a fold change of 1.65; $P < 0.001$, whereas the most significant hypermethylated region was found in the HOXD3 (Figure 1D) promoter, a gene which we have previously validated in an independent cohort with methylation strongly correlated with Gleason score (23, 29). In the GS7 versus GS8 analysis, 61 probes representing 50 regions were differentially methylated (Supplementary Table S1). Contrary to what was found in GS6 versus GS7, the majority of the regions were hypermethylated (30) as compared with hypomethylated (12) in GS8 as compared with GS7. Hypermethylation of the PITPNM1 promoter was the most significant of all DMRs (fold change $= 1.56; P < 0.001$), whereas the most significant hypomethylated region was intragenic to BCAT2 (fold change $= 1.56; P = 0.001$). Among the genes with greater methylation in GS $\geq 8$ versus GS7 was TBX15, another gene which we have previously published as correlated with Gleason score (23).

As the Agilent CpG island platform had representation within promoter, intragenic, and intergenic regions, we tested the distribution of hypomethylated versus hypermethylated regions in the total Gleason score analysis. The distribution of probe locations was significantly different between the two, with a greater fraction of hypermethylated probes occurring in intergenic regions and smaller fraction occurring in promoter regions (Fig. 1A; $\chi^2 P = 5.79 \times 10^{-5}$).

**ERG-related DNA methylation changes**

Next, we sought to determine if there were gene-specific DNA methylation differences by stratifying the microarray cases based on ERG protein expression status. Of the 39 prostate cancer cases submitted for CpG island array analysis, 16 were positive for ERG expression, 18 were negative, and 5 were indeterminate due to insufficient tumor material. The summary of ERG expression status along with clinicopathologic characteristics of this cohort is in Supplementary Table S2.

The comparison of ERG-positive and -negative cases revealed a total of 193 differentially methylated probes representing 123 unique locations, with 100 hypermethylated and regions only 23 hypomethylated regions (Supplementary Table S3).
Table 1. Significant genomic regions showing hypermethylation correlating with Gleason score in ERG-positive cases and hypomethylation correlating with Gleason score in ERG-negative cases

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Probe coordinates</th>
<th>ERG negative</th>
<th>ERG positive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GS7-10 vs. GS6</td>
<td>ANOVA</td>
</tr>
<tr>
<td>chr5:002696343-002696402</td>
<td>chr5:002696343-002696402</td>
<td>-1.607095846</td>
<td>0.013</td>
</tr>
<tr>
<td>N4AA2</td>
<td>chr2:1568885322-156885371</td>
<td>-1.801580198</td>
<td>0.017</td>
</tr>
<tr>
<td>ICMT-HES3</td>
<td>chr1:006224560-006224619</td>
<td>-1.408132896</td>
<td>0.036</td>
</tr>
<tr>
<td>TBX15</td>
<td>chr1:161981774-161981818</td>
<td>-1.297381807</td>
<td>0.033</td>
</tr>
<tr>
<td>PRKCE</td>
<td>chr2:045732845-045732889</td>
<td>-1.25398302</td>
<td>0.036</td>
</tr>
<tr>
<td>chr13:111675469-111675518</td>
<td>chr13:111675469-111675518</td>
<td>-1.261257411</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table S3). Intragene hypermethylation of the PRDM16 gene was the most significant DMR (Figure 2D; fold change = 2.33; P < 0.001), whereas the most significant hypomethylated region was in the CREM 1 promoter (Figure 2C; fold change = 2.71; P < 0.001). Of note, TBX15 methylation was also significantly associated with ERG expression, whereas HOXD3 met the threshold criteria for P value (4 probes < 0.05) but did not meet the fold change criteria (maximum fold change = 1.79). CYP26AI was not significant either by fold change or P value. This suggests that our statistical analysis suitably detects strong associations but may miss weaker associations, as correlations observed in our previously published larger validation cohort were strongest for TBX15, followed by HOXD3 and finally CYP26AI (23).

Furthermore, of the 3 overlapping regions between our arrays and the top 100 DMRs published by Borno and colleagues, 2 were consistent (CACNA1D and GREM1), whereas the third (FOXE3) was not detected in our array analysis, providing further evidence that our DMRs were specific but may lack some sensitivity.

We then tested if any of the differentially methylated probes discovered in the Gleason score analysis were similarly differentially methylated according to ERG expression status. Only 7 DMRs were overlapping between the 2 analyses (Figure 2B and Table 1), suggesting that ERG-related methylation events are mostly not indicative of Gleason score-related methylation changes.

Similar to the Gleason score hypermethylated probes, approximately 50% of ERG-related hypermethylated locations were found in intergenic regions, whereas there were fewer hypermethylation events in gene promoter regions. All together, the genomic locations of hypermethylated versus hypomethylated probes were significantly different (Fig. 2A; P = 0.001).

All of the above results were based on stratification of cases into distinct categories based on Gleason score or ERG expression. Therefore, to group samples together based on methylation profiles in an unbiased fashion, we conducted hierarchical clustering of cases and probes using the CpG island microarray data. From this analysis, 193 of the most valuable probes (≥1.5 SD across all 39 prostate cancer specimens) representing 110 unique genes distinguished 2 distinct clusters of cases, with 14 cases in cluster 1 and 25 cases in cluster 2 (Fig. 3). Interestingly, ERG expression status was significantly different between clusters 1 and 2, with 1 of 10 cases in cluster 1 ERG-positive and 17 of 24 cases in cluster 2 ERG-positive (Fisher exact P = 0.002).

Gleason score and ERG

As ERG expression status seemed to have an effect on the methylation level of numerous genes and thus may confound overall Gleason score results, we next tested for differences between Gleason score within each ERG subtype (positive or negative). In the ERG-negative group, we discovered 412 differentially methylated probes representing 388 genomic regions (Supplementary Table S3) that were associated with Gleason score, with the overwhelming majority being hypomethylated (380 of 388; 97.9%). The majority of the hypomethylated regions were localized to promoters (Figure 4A, 50.8%), followed by intergenic and intragenic locations (30.2% and 19.1%, respectively).

The trend of more hypomethylated regions was also present in the ERG-positive group, although not to the same extent. Of the 580 genomic regions differentially methylated between Gleason score in ERG-positive prostate cancer, 374 were hypomethylated (64.5%), whereas 206 (35.5%) were hypermethylated (Supplementary Table S3). Again, the differences between hypomethylated and hypermethylated locations were striking, with 64.2% of hypomethylated regions existing in gene promoters compared with 22.3% for hypermethylated regions. Conversely, intergenic locations comprised 22.5% of the hypomethylated regions, whereas 48.1% of hypermethylated regions were intergenic (P = 4.25 × 10^-15).

Finally, we compared DMRs that associate with Gleason score in the overall Gleason score analysis and the ERG-stratified Gleason score analysis (Fig. 4B–D). Of the 388 regions associated with Gleason score in ERG-negative cases, 98 overlapped with regions significant in the total Gleason score analysis that did not take ERG expression into consideration. Similarly, 87 regions significantly associated with Gleason score in ERG-positive cases overlapped with probes significant in the total Gleason score analysis. By removing such overlapping probes, we found 487 gene
regions associated with Gleason score unique to ERG-positive cases and 284 gene regions associated with Gleason score unique to ERG-negative cases. Interestingly, we noted 6 unique instances where hypermethylation of a specific probe was associated with Gleason score in ERG-positive specimens, whereas hypomethylation of the same region was associated with Gleason score in ERG-negative specimens (Table 1). There were no instances where hypomethylation was associated with Gleason score in ERG-positive cancers, whereas hypermethylation of the same region was associated with Gleason score in ERG-negative prostate cancer.

**Functional annotation**

To identify groups of genes belonging to different biologic pathways that may be targeted by DNA methylation, we conducted DAVID functional clustering analysis using the regions detected from analysis of Gleason score and ERG stratification data. Using the functional annotation enrichment feature, there was strong enrichment for DNA-binding proteins/transcription factors in all datasets (Supplementary Tables S4 and S5). In particular, homeobox gene hypermethylation was strongly enriched in higher-grade prostate cancer as were genes involved in neuronal differentiation (Benjamini–Hochberg–adjusted $P = 1.91 \times 10^{-10}$ and $4.2 \times 10^{-9}$, respectively). For hypomethylated regions, genes involved in RNA metabolism were significantly enriched as were DNA-binding proteins (Benjamini–Hochberg–adjusted $P = 0.002$ and $2.11 \times 10^{-5}$, respectively). With respect to hypermethylated DMRs based on ERG expression, homeobox genes were again strongly enriched as were genes involved in positively regulating transcription from RNA pol II promoters (Benjamini–Hochberg–adjusted $P = 1.20 \times 10^{-9}$ and .002, respectively). Similar strong enrichment of homeobox genes was found in Gleason score–associated hypermethylated regions in ERG-positive cases only (Supplementary Table S5). Not...
included in this analysis were ERG-associated hypomethylated regions or Gleason score–associated hypermethylated regions in ERG-negative cases, as there were an insufficient number of regions/genes to conduct a meaningful analysis (N < 25).

**DNA methylation and polycomb-repressive marks**

As all hypermethylated datasets (Gleason score and ERG associated) had strong enrichment of homeobox transcription factors and these genes are well-characterized targets of PRCs, we assessed the overlap between hypermethylated

---

**Table:**

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS</td>
<td>P = 0.891</td>
</tr>
<tr>
<td>ERG Status</td>
<td>P = 0.002</td>
</tr>
</tbody>
</table>

**Figure 3.** Unsupervised hierarchical clustering of prostate cancer specimens. Clustering of all cases and the 193 probes with SD ≥ 1.5 across all prostate cancer specimens. A, Pearson χ² analysis of Gleason score distribution separated by clusters 1 and 2; B, Fisher exact analysis separated by clusters 1 and 2.
DMRs associated with Gleason score and/or ERG with publicly available data analyzing the PRC associated histone modification H3K27me3 in benign RWPE-1 cells (28). ChIP of H3K4me3, H3Ac, H3K9me3, and RNA pol II in RWPE-1 and methylated DNA immunoprecipitation of RWPE-1, PrEC, and 5 prostate cancer cell lines was also conducted in this dataset and used in this analysis. In the Gleason score–associated hypermethylation probeset, there was significantly greater H3K27me3 (Fig. 5A; \( t \) test; \( P = 1.68 \times 10^{-10} \)) along with fewer than expected regions with active histone marks H3Ac, H3K4me3 as well as RNA pol II binding (\( t \) test; \( P = 8.00 \times 10^{-17} \), and \( 1.77 \times 10^{-12} \), and \( 3.38 \times 10^{-11} \), respectively), suggesting that many of the hypermethylated DMRs are in a transcriptionally silent state in benign cells.

In the ERG-positive hypermethylated and ERG-positive hypermethylated associated with Gleason score datasets, we again observed enrichment of H3K27me3 in RWPE-1 cells (\( t \) test; \( P = 2.62 \times 10^{-10} \) and \( 6.44 \times 10^{-17} \), respectively), along with depletion of RNA pol II binding (\( t \) test; \( P = 8.79 \times 10^{-26} \) and \( 7.71 \times 10^{-35} \), respectively).

To confirm that enriched H3K27me3 was not a unique feature of HPV-18–immortalized RWPE-1 cells, we used data generated from benign PrEC cells on a separate platform (19). This dataset analyzed DNA methylation, H3K27me3, and SUZ12 binding in both PrEC and PC3 cells. There was strong enrichment of H3K27me3 in both PrEC and PC3 cells within the 17 regions from this dataset, which overlapped with our 106 Gleason score hypermethylated DMRs (Fig. 5B; \( t \) test; \( P = 9.23 \times 10^{-18} \) and \( 3.28 \times 10^{-5} \), respectively). SUZ12 had enriched binding in PrEC cells, but not in PC3 cells, for regions localized to our Gleason score–associated hypermethylated DMRs (\( t \) test; \( P = 2.25 \times 10^{-9} \) and 0.91, respectively), suggesting that while some H3K27me3 polycomb-repressive marks remain, PRC2 binding is lost at these loci in PC3 cells. Not surprisingly, we observed increased levels of DNA methylation for our Gleason score hypermethylated DMRs in PC3 cells.
cells, whereas PrEC cells were significantly depleted of methylation ($t$ test; $P = 4.28 \times 10^{-17}$ and $2.33 \times 10^{-5}$, respectively).

In addition, we analyzed prostate cancer prognostic DNA methylation markers discovered by Mahapatra and colleagues (18) in conjunction with the ChIP-chip datasets to assess if our findings were unique to our methylation dataset and/or platform. We found a significantly greater level of H3K27me3 in both ChIP-chip datasets (Fig. 5A and B; $P = 4.24 \times 10^{-13}$ and $2.96 \times 10^{-15}$) and a reduced level of RNA pol II binding in benign cells (Fig. 5A; $P = 4.23 \times 10^{-3}$), confirming the findings from our own methylation data.

**Discussion**

Prostate cancer is a clinically and morphologically heterogeneous disease and Gleason score remains one of the best markers of prostate cancer prognosis, whereas the TMPRSS2:ERG gene fusion is one of the most common genetic aberrations found in prostate cancer, which conveys no independent prognostic significance. PCR expression and epigenetic signatures are also being realized as indicators of aggressive cancers and specific cancer subtypes (31, 32). We report here that the DNA methylation landscape in prostate cancer is altered with respect to Gleason score and expression of the ERG oncogene, and that unique sets of differentially methylated loci correlate with Gleason score when prostate cancer is stratified into ERG-negative and -positive subtypes. Furthermore, we report that the majority of hypermethylation events correlating with Gleason score and ERG occur in genes with polycomb-repressive marks in benign prostate epithelial cells.

We have previously reported DNA methylation changes that occur between high grade, GS $\geq$ 8, and low grade, GS6 prostate cancer (24). In this study, we have included intermediate grade GS7 to assess changes that occur across the entire Gleason score spectrum and that occur specifically from GS6 to GS7 or GS7 to GS $\geq$ 8. GS7 tumors vary markedly in disease course with some following a more aggressive course and others a more indolent one. Thus, analysis of GS7 methylomes may better separate indolent versus aggressive tumors. We found that many of the DMRs are more prevalent during the transition from GS6 to GS7 or from GS7 to GS $\geq$ 8. For example, methylation of the SESN3 promoter was significantly greater in the GS7 versus GS $\geq$ 8 analysis (fold change 1.53; $P = 0.006$) and not in GS6 versus GS7 (fold change = 1.18; $P = 0.461$). It is important to note, however, that although many DMRs were only significant in either the GS6 versus GS7 or GS7 versus GS $\geq$ 8 analysis, the direction of change in the nonsignificant comparison remained consistent (either increased or decreased methylation). The direction of change for each significant region was also consistent in the total Gleason score analysis. These results may be explained by the Gleason patterns present within each Gleason score. GS6 specimens in this study consisted of pure grade 3 pattern, whereas GS7 was by definition a mixture of grade 3 and 4 patterns and GS $\geq$ 8 was pure grade 4 pattern. Therefore, differential methylation might be expected that is dependent on the proportion of grade 3 and 4 within the sample. We have previously reported such correlations with Gleason pattern (23, 29, 33).

Using genome-wide CpG island DNA methylation profiling, we report 123 unique DMRs that correlate with ERG oncogene expression. Of note, CACNA1D and HEXA hypomethylation was found in ERG-positive cases compared with ERG-negative. Both of these genes have been described as potential ERG target genes with increased expression in ERG-positive cancers (34, 35). In addition, hierarchical clustering of the top 193 variable probes from the microarray dataset revealed one methylome cluster with 17 of 24 cases positive for ERG expression, with the other cluster containing 1 of 10 cases positive for ERG expression. This observed relationship between ERG expression and gene specific DNA methylation is not surprising given the role of ERG as a transcription factor that likely requires epigenetic cofactors in modifying and remodeling surrounding chromatin to exert its effects. Recent functional work in ERG-overexpressing cells has identified global changes in chromatin interactions that are likely dependent on chromatin remodeling and...
DNA Methylation, ERG Expression, and Polycomb Target Genes

The functional consequence of hypermethylated DMRs localized to homeobox genes and regions of H3K27me3 domains remains unknown. As homeobox genes are master regulators, one possibility is that DNA methylation disturbs PRC-mediated regulation and may create a more stable silent state as opposed to a more dynamic state. Histone methylation, including H3K27me3, can be removed via well-known histone demethylases (47), whereas active DNA demethylation seems to be a more cumbersome process involving hydroxymethylation and DNA repair pathways (48, 49). Another possibility is that genes with PRC marks in benign prostate cells are particularly susceptible to DNA methylation in prostate cancer, but that the latter is merely a by-product of overactive DNA methylation machinery.

In conclusion, we present genome-wide DNA methylation signatures correlating with prostate cancer Gleason score, with ERG oncogene expression status, and with Gleason score stratified by ERG expression status. Furthermore, we show that in each of these analyses, DNA hypermethylation preferentially occurs at CpG islands with polycomb-repressive marks in benign prostate epithelial cells. Currently, this has unknown functional consequences, but may be useful in establishing patient prognosis by assessing epigenetic changes at polycomb-repressed loci. In addition, integrating ERG status into methylation biomarker discovery expands the list of potential events associated with poor prostate cancer prognosis, in theory by removing the effect of a potential confounding variable in prostate cancer biomarker studies. Thus, the epigenetic landscape of prostate cancer is dependent on the presence or absence of this fusion and can give clues about disease progression in prostate cancer.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: K. Kron, V. Pethe, T. van der Kwast, B. Bapat
Development of methodology: K. Kron, D. Trudel, V. Pethe, B. Bapat
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D. Trudel, N. Fleshner, T. van der Kwast, B. Bapat
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): K. Kron, L. Briollais, B. Bapat
Writing, review, and/or revision of the manuscript: K. Kron, D. Trudel, V. Pethe, N. Fleshner, T. van der Kwast, B. Bapat
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): B. Bapat
Study supervision: T. van der Kwast, B. Bapat

Grant Support
This work was supported by the National Cancer Institute of Canada/Canadian Prostate Cancer Research Initiative #18568 [to B. Bapat, T. van der Kwast, and N. Fleshner], the Ontario Student Opportunity Trust Fund [to K. Kron], the Ontario Graduate Scholarships [to K. Kron], and The Terry Fox Foundation Strategic Health Research Training Program in Cancer Research at Canadian Institute of Health Research (CIHR) and Ontario Institute of Cancer Research [to D. Trudel].

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 October 4, 2012; revised February 28, 2013; accepted March 7, 2013; published OnlineFirst April 2, 2013.
References


Cronk et al.


Altered DNA Methylation Landscapes of Polycomb-Repressed Loci Are Associated with Prostate Cancer Progression and ERG Oncogene Expression in Prostate Cancer

Ken Kron, Dominique Trudel, Vaijayanti Pethe, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-12-3139

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2013/03/29/1078-0432.CCR-12-3139.DC1

Cited articles
This article cites 48 articles, 17 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/19/13/3450.full.html#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.