DNA Hypomethylation Arises Later in Prostate Cancer Progression than CpG Island Hypermethylation and Contributes to Metastatic Tumor Heterogeneity

Srinivasan Yegnasubramanian,1 Michael C. Haffner,1,7 Yonggang Zhang,1,2 Bora Gurel,3 Toby C. Cornish,5 Zhijin Wu,7 Rafael A. Irizarry,5 James Morgan,1,3 Jessica Hicks,1,3 Theodore L. DeWeese,7 William B. Isaacs,1,7 G. Steven Bova,1,7 Angelo M. De Marzo,1,3,4 and William G. Nelson1,2,3,4

1Sidney Kimmel Comprehensive Cancer Center, 2Department of Radiation Oncology, 3Department of Pathology, 4Brady Urological Institute, 5Department of Biostatistics, School of Public Health, Johns Hopkins University, Baltimore, Maryland; 6Center for Statistical Sciences, Brown University, Providence, Rhode Island; and 7Innsbruck Medical University, Innsbruck, Austria

Abstract
Hypomethylation of CpG dinucleotides in genomic DNA was one of the first somatic epigenetic alterations discovered in human cancers. DNA hypomethylation is postulated to occur very early in almost all human cancers, perhaps facilitating genetic instability and cancer initiation and progression. We therefore examined the nature, extent, and timing of DNA hypomethylation changes in human prostate cancer. Contrary to the prevailing view that global DNA hypomethylation changes occur extremely early in all human cancers, we show that reductions in 5mC content in the genome occur very late in prostate cancer progression, appearing at a significant extent only at the stage of metastatic disease. Furthermore, we found that, whereas some LINE1 promoter hypomethylation does occur in primary prostate cancers compared with normal tissues, this LINE1 hypomethylation is significantly more pronounced in metastatic prostate cancer. Next, we carried out a tiered gene expression microarray and bisulfite genomic sequencing–based approach to identify genes that are silenced by CpG island methylation in normal prostate cells but become overexpressed in prostate cancer cells as a result of CpG island hypomethylation. Through this analysis, we show that a class of cancer testis antigen genes undergoes CpG island hypomethylation and overexpression in primary prostate cancers, but more so in metastatic prostate cancers. Finally, we show that DNA hypomethylation patterns are quite heterogeneous across different metastatic sites within the same patients. These findings provide evidence that DNA hypomethylation changes occur later in prostate carcinogenesis than the CpG island hypermethylation changes and occur heterogeneously during prostate cancer progression and metastatic dissemination. [Cancer Res 2008;68(21):8954–67]

Introduction
Aberrations in genomic DNA methylation occur almost universally in human cancers. However, the pattern of DNA methylation abnormalities in cancer cells seems paradoxical: compared with normal cells, cancer cells seemed concomitantly hypermethylated at specific CpG island sequences but hypomethylated at CpGs found in most other sites, resulting in a net loss of genomic 5mC content (1–3). Hypermethylation of promoter CpG islands can silence important tumor-suppressor genes and other genome protection genes (4). Hypomethylation of genomic DNA has been associated with increased rates of genomic instability (5, 6), perhaps due to decondensation of chromatin into recombination-permissive conformations (7) or due to the activation of retrotransposon elements (8). A few recent studies have linked genomic hypomethylation with increased genetic instability in human cancers (9–12). Another possibility is that CpG island sequences at genes that are normally methylated and silenced in a tissue-specific manner (13, 14) may become aberrantly hypomethylated, causing inappropriate activation of the corresponding gene in cancer cells. Such aberrations in lineage-specific marks could directly parallel the lack of appropriate differentiation and morphology often seen in cancer cells.

Understanding the relative timing of DNA hypermethylation and hypomethylation changes in cancer is crucial to understanding the importance of these changes in various stages of disease initiation and progression. One prevailing model of epigenetic dysregulation in carcinogenesis and disease progression has held that, on average, reduction of DNA cytosine methylation seems to occur early in carcinogenesis in human cancers at the stage of premalignant lesions or even in aging benign tissues and that this hypomethylation precedes and perhaps sets the stage for later gene promoter CpG island hypermethylation changes (reviewed in ref. 15). Some have even hypothesized that age-related decreases in DNA methylation may provide a “field-effect” for cancer initiation, perhaps partially explaining the association of most malignancies with advancing age (9, 15).

However, there are hints that prostate cancer may not adhere to this generalized model. An early report showed that reduction in overall 5mC content was rare in primary prostate cancers but frequent in metastatic prostate cancer (16). Additionally, LINE1 hypomethylation seemed to occur after hypermethylation of CpG island sequences in the 5’ upstream regions of key genes (17). However, these studies examined limited numbers of prostate specimens, did not analyze multiple stages of prostate cancer.
progression including metastases, or did not thoroughly examine multiple forms of DNA hypomethylation changes. Here, we perform the most comprehensive examination to date of the nature, extent, and timing of DNA hypomethylation changes in prostate cancer disease progression. Specifically, we hypothesize that such DNA hypomethylation changes can manifest in at least three major ways: (a) reduction in genomic 5meC content, (b) decreased methylation of repetitive sequences, such as LINE1 elements, and (c) hypomethylation of normally methylated promoter CpG island sequences, leading to overexpression of the associated genes. We show that these hypomethylation changes occurred late in prostate cancer disease progression, manifesting most prominently at the stage of metastatic disease. Furthermore, compared with CpG island hypermethylation changes, DNA hypomethylation changes were heterogeneous across different metastatic deposits within the same cases. Taken together, these data suggest that, contrary to the prevailing model of epigenetic dysregulation, DNA hypomethylation occurs late in prostate cancer disease progression and is unlikely to be involved in prostate cancer initiation and more likely to be involved in the formation and propagation of metastases.

**Materials and Methods**

**Cell culture, DNA, and RNA extraction.** Primary cultures of normal prostate epithelial cells (PrEC), two prostate stromal cell models (4ST and 1S), and six prostate cancer cell lines (LNCaP, C4-2B, PC-3, DU-145, LAPC-4, CWR22Rv1) were propagated in tissue culture as described previously (18). Cells were allowed to reach 80% confluence in tissue culture flasks. Whole RNA and genomic DNA were isolated using the RNeasy kit (Qiagen) and the DNeasy Tissue kit (Qiagen), respectively, by following the manufacturer’s protocols. RNA and DNA were quantitated by 260-nm UV absorbance using a standard spectrophotometer.

**Tissue specimens, tissue microarrays, and DNA isolation.** Tissues from primary prostate cancer tissues from 76 men undergoing radical prostatectomy, tumor-adjacent benign tissues from 12 of these men, multiple anatomic compartments of prostate cancer, lymph node metastases from 8 men undergoing radical prostatectomy for clinically presumed localized prostate cancer (the radical prostatectomy was aborted in these men due to the discovery of metastatic lymph nodes), and normal prostate specimens from 13 brain dead organ donors with no evidence of prostate disease were obtained and subjected to DNA isolation, as previously described (18). The tumor-adjacent benign prostate tissues included seven specimens containing extensive regions of PIN, five specimens containing extensive regions of PIA, and one specimen with neither PIA nor PIN. See Supplementary Table S1 for a detailed description of evaluation of DNA specimens available from tissue specimens and average age and age range of each DNA specimen group. Two hundred and twenty core tissue microarrays were constructed from formalin-fixed paraffin-embedded tissue blocks as previously described (19). All studies with human specimens were approved by our institutional review board.

**Quantitation of genomic 5meC content by high-pressure liquid chromatography-tandem mass spectrometry.** The overall 5meC (5-methyl-2-deoxycytidine) content [as percentage of total 2-deoxycytidine (2dC) content] in genomic and control DNAs was determined by a high-pressure liquid chromatography-tandem mass spectrometry (LC-MS/MS) procedure as described elsewhere (20, 21), with minor modifications. Please see supplementary materials for a detailed description.

**Gene expression microarray analysis, bioinformatics, and statistical analysis of methylation data.** Ten micrograms of total RNA from PrEC, 4ST, 1S, LNCaP PC-3, DU-145, LAPC-4, C4-2B, and CWR22Rv1 cells were processed, labeled, and hybridized to hgu133A whole genome gene expression microarrays (Affymetrix) according to the manufacturer’s protocols. The Bar Code gene expression microarray analysis (22), with minor modifications, was used to identify genes that were expressed in the cancer cell lines with high confidence but not expressed in the PrEC, 4ST, and 1S normal cells. Please see supplementary materials for a detailed description of microarray analysis procedures, bioinformatics analysis of the distribution of CpG dinucleotides in the human genome, and statistical methods used to analyze methylation data.

**Bisulfite genomic sequencing.** Bisulfite genomic sequencing was carried out on genomic DNA samples that were bisulfite converted using the EZ DNA methylation kit (Zymo Research) as previously described (23). Primers and annealing temperatures used to amplify bisulfite converted genomic DNA are listed in Supplementary Table S3.

**COMPARE methylation assay.** A variation of the previously described COMPARE-MS assay (23), termed the COMPARE (combination of methylated DNA precipitation and restriction enzyme digestion) assay, was used to determine the degree of CpG island methylation at candidate hypomethylated genes and at LINE1 consensus promoter sequences in the prostate cell line and tissue samples. The COMPARE assay is identical to the COMPARE-MS assay except that the methylation-sensitive restriction enzyme was omitted (23). A detailed description of the COMPARE assay can be found in supplementary materials, with PCR primers listed in Supplementary Table S4.

**Immunohistochemistry and analysis.** Immunohistochemical staining was carried out on deparaffinized and rehydrated tissue sections from formalin-fixed, paraffin-embedded blocks or tissue microarrays (TMA). Immunohistochemistry for NY-ESO-1 (encoded by CTAGIB gene) and MAGEA1 was done with the EnVision+ kit (DakoCytomation). For NY-ESO-1, slides were dewaxed for 10 min in xylene, rehydrated in graded ethanol, and steamed for 15 min in 1× EDTA buffer (Zymed) for 40 min and stained with anti-NY-ESO-1 primary antibody (1:100 dilution; Zymed) at room temperature for 45 min. Immunohistochemistry with anti-5meC antibody was done with the Power Vision+ detection system (Immunovision Technologies). Antigen retrieval was accomplished by steaming slides for 20 min in citrate unmasking solution (Vector Labs) for antigen retrieval and then incubated with anti–NY-ESO-1 primary antibodies (1:100 dilution; Zymed) at room temperature for 45 min. Immunohistochemistry with anti-5meC antibody was done with the Power Vision+ detection system (Immunovision Technologies). Antigen retrieval was accomplished by steaming slides for 20 min in citrate unmasking solution and then treating with 3.5 N HCl for 15 min. Slides were then stained with anti-5meC primary antibody (1:4,000 dilution; Calbiochem). The TMA slides were digitally scanned, and the immunohistochemical staining was analyzed using the FrIDA image analysis tool (24). See supplementary materials for a detailed description of evaluation of immunohistochemical staining and quantitative image analysis using FrIDA.

**Results and Discussion**

**Genomic 5meC content is reduced in metastatic but not primary prostate cancers.** DNA methylation occurs predominantly in CpG dinucleotides in adult somatic tissues in vertebrate genomes. Because CpG dinucleotides are nonuniformly distributed in the human genome (25–27), we examined their distribution in the general genomic 5meC content as a fraction of total genomic 2dC content (28, 29). Approximately 4.85% of all cytosines occur in CpG dinucleotides in the human male genome. Of these CpGs, ~23% occur in CpG islands, in which one fifth to one fourth of cytosines occur in the context of CpG dinucleotides. The remaining ~77% of CpGs occur very sparsely throughout the genome with approximately one twentieth of the cytosines occurring in CpG sequences. Therefore, measurements of overall genomic 5meC content most likely reflect the degree of methylation of these sparsely spaced CpG dinucleotides in the human genome.

Using a highly sensitive LC-MS/MS approach, we measured overall genomic 5meC content as a fraction of total genomic 2dC content (5meC + 2dC) in a collection of prostate cell lines and control specimens (Fig. 1A). This method was capable of accurately detecting 2 fmol of 5meC and 2dC, showed linearity over a broad
dynamic range spanning two to three orders of magnitude, and allowed robust measurement of overall 5mC content from as little as 1 ng of genomic DNA (data not shown). This method was highly accurate when tested with control oligonucleotides containing 0% (UM-Oligo; data not shown) and 80% (M-Oligo 80) actual 5mC content.

Furthermore, the M.SssI-treated WBC genomic DNA, in which all CpG dinucleotides were enzymatically methylated in vitro, had a
5meC content of ~4.9%, representing the maximal amount of CpG methylation possible in a normal human genomic DNA specimen. Note that this is also an estimate of the percentage of cytosines in CpG dinucleotides in the human genome and is in close agreement with the in silico analysis of CpG content in the human genome, validating this assay for analysis of genomic DNA samples. Untreated normal WBCs and normal PrEC and prostate stromal (4ST) cells had 5meC contents between ~3.5% and 4.0%, suggesting that ~70 to 80% of all CpGs were methylated in these nonmalignant cells. In comparison to these normal cells, five of the six prostate cancer cell lines (LNCaP, C4-2B, LAPC-4, CWR22Rv1, and PC-3) had significantly reduced 5meC content (2.0–2.9%) with only DU-145 showing a 5meC content (3.9%) that was not reduced.

We next examined genomic DNA from a large collection of prostate tissues (Fig. 1B). The 5meC content of benign prostate tissues resembled that of the WBC, PrEC, and 4ST cells. Surprisingly, compared with benign prostate specimens, overall

Figure 2. Compared with benign prostate, reduction of LINE1 promoter methylation occurs to some extent in primary prostate cancer but to a greater extent in metastatic prostate cancer. A, an overview of the COMPARE assay used to detect LINE1 hypomethylation and gene-specific CpG island hypomethylation. Each sample is split into an untreated fraction and an M.SssI-treated fraction, in which all CpG dinucleotides are methylated to completion using the M.SssI methyltransferase enzyme. Each fraction is fragmented with restriction enzymes, and then densely methylated fragments are captured and enriched with magnetic bead immobilized MBD2-MBD polypeptides that specifically bind to densely methylated genomic DNA fragments with high avidity. Enriched fragments are eluted, and the amount of a given genomic region of interest is determined by real-time PCR. The ratio of the amount of methylated fragments in the untreated fraction to that in the M.SssI-treated sample, called MI, gives an estimate of the fraction of input genomic DNA fragments that were methylated at a given region of interest. B, LINE1 promoter methylation is reduced to varying extents in prostate cancer cell lines compared with normal PrEC and stromal (4ST) cells and normal WBCs as seen by application of the COMPARE assay for detection of methylation at consensus LINE1 promoter regions. C, application of the COMPARE LINE1 promoter hypomethylation assays to genomic DNA from prostate tissues shows that primary prostate cancers (red) and metastatic prostate cancers (brown and maroon) have significantly reduced LINE1 promoter methylation compared with organ donor normal prostates (green) and tumor-adjacent benign prostate tissues, including PIA and PIN lesions (green). Box-and-whisker plots of MI at LINE1 promoter elements. The extent of LINE1 promoter methylation was even more pronounced in metastatic prostate cancers compared with primary prostate cancers. Number of specimens analyzed for each prostate tissue category is indicated. * for the autopsy metastases, 76 metastatic deposits from 28 subjects (one to seven metastatic deposits from each subject) were analyzed. Results of t test for all statistically significant differences in the mean 5meC content between all pairwise combinations are shown. Differences that remained significant after adjusting for age are indicated as follows: *, P < 0.05; †, P < 0.005.
Figure 3. Identification of genes that are not expressed in normal PrEC and stromal cells but expressed with high confidence in prostate cancer cell lines. Hierarchical clustering and heat mapping of Z scores from Bar Code analysis of gene expression data from Affymetrix hgU133a microarray experiments show that this analysis identified 168 genes with absence of significant expression in the normal PrEC and stromal cells but very high confidence of absolute expression in at least one of the prostate cancer cell lines. These genes were considered candidate hypomethylated genes. Green bars, known cancer-testis antigen genes identified by this analysis; red bars, genes that were selected for bisulfite sequencing analysis. Details of all genes in this figure are given in Supplementary Table S1.
DNA Hypomethylation in Prostate Cancer Progression

5mC content was reduced in the metastatic (P < 0.0001) but not primary prostate cancers or tumor-adjacent PIA/PIN/normal tissues. The 5mC content of these metastases resembled that of the prostate cancer cell lines, which were originally derived from prostate cancer metastases (28–31). Furthermore, there was a statistically significant trend in 5mC content with primary cancer > hormone naive lymph node metastases > hormone refractory autopsy metastases (P_{trend} < 0.0001). Immunohistochemical analysis and quantitative image analysis of anti-5mC antibody–stained TMAs containing 220 cores with an assortment of organ donor normal prostate tissues, primary prostate cancer tissues from radical prostatectomy, and metastatic prostate cancers from a rapid autopsy series confirmed that reductions in 5mC content occurred in metastatic but not primary prostate cancer tissues (Fig. 1C and D). Taken together, these findings suggest that overall decreases in genomic 5mC content occur late in prostate cancer progression at the stage of metastatic disease. This is in stark contrast with the conventional view that reductions in overall 5mC content generally occur early in carcinogenesis and disease progression.

Degree of LINE1 promoter methylation is reduced in metastatic prostate cancer compared with primary prostate cancer and benign prostate. Repetitive elements constitute nearly half of all sequences and half of the CpG content in the human genome (Supplementary Fig. S1B; refs. 26, 27). LINE1 (L1) elements are the largest class of repetitive elements in the human genome, constituting ~16% of all sequences and ~9.5% of all CpGs. Of particular interest, LINE1 elements possess an internal promoter, which contains a high density of CpG dinucleotides that can meet criteria for a CpG island (32). The LINE1 promoter CpG island is typically methylated (33), and its hypomethylation can lead to transcriptional activation of intact LINE1 elements, induce retrotransposition, and facilitate genetic instability (34, 35). The vast majority of LINE1 elements, however, are truncated or highly degenerate, lack the promoter sequence, and are not capable of active retrotransposition (26, 36). Because they constitute a small fraction of all CpGs in the human genome, LINE1 promoters could become hypomethylated even without significant overall reductions in genomic 5mC content. Therefore, we investigated the possibility that LINE1 promoter CpG islands may become hypomethylated during the initiation and progression of prostate cancer.

We used a derivative of the recently described COMPARE-MS assay (23), which takes advantage of the high affinity and selectivity of the methyl-binding domain fragment of MBD2 (MBD2-MBD) in binding symmetrically methylated DNA, to assess the extent of LINE1 methylation in prostate specimens. In the modified assay (Fig. 2A), which we term COMPARE, each genomic DNA specimen was split into an untreated fraction and an M.SssI-treated fraction, in which all CpGs were fully methylated. These samples were fragmented by restriction enzyme digestion and subjected to enrichment for methylated DNA fragments by capture with MBD2-MBD polypeptides immobilized on magnetic beads. Enriched methylated LINE1 promoter fragments were then detected by real-time PCR for the consensus LINE1 promoter (32). The methylation index (MI), which is the ratio of the signal from the untreated fraction to the M.SssI-treated fraction for each specimen, gives a measure of the percentage of input copies that were methylated at a given locus of interest. In comparison with the original assay, omission of the methylation-sensitive restriction enzyme in this COMPARE assay allowed more stringent quantitation of hypomethylation because only sequences that are sufficiently undermethylated to escape capture with the MBD2-MBD would seem hypomethylated. The COMPARE assay exhibited a linear dynamic range spanning more than three orders of magnitude and could be used to accurately measure DNA methylation from as little as 32 pg of genomic DNA (23). Furthermore, the COMPARE LINE1 promoter hypomethylation assay showed strong concordance with results of bisulfite genomic sequencing of LINE1 promoters in reference samples, showing that this assay had a linear dynamic range from no LINE1 methylation to complete LINE1 methylation (Supplementary Fig. S7).

Application of the COMPARE LINE1 hypomethylation assay revealed that prostate cancer cell lines had less LINE1 promoter methylation than the normal WBC, PrEC, and 4ST cells, which showed ~80% to 90% methylation of input LINE1 promoter sequences. In particular, the LNCaP and C4-2B cell lines had less than half the LINE1 promoter methylation levels as the normal cells (Fig. 2B). The benign prostate tissues from organ donors and the PIA/PIN/tumor-adjacent benign prostate tissues exhibited high levels of LINE1 methylation similar to the WBC, PrEC, and 4ST genomic DNA specimens (Fig. 2C). In comparison with these benign specimens, the primary prostate cancers had decreased LINE1 methylation, although they did not exhibit a significant reduction in overall genomic 5mC content. This finding is in agreement with previous reports showing decreased LINE1 element methylation in primary prostate cancers (17). It is possible that DNA hypermethylation changes in the primary prostate cancers sufficiently compensate for the decreased LINE1 methylation so that there is no net reduction in the genomic 5mC content. Finally, metastatic prostate cancers had even lower LINE1 methylation than the primary prostate cancers. For LINE1 promoter methylation also, there was a statistically significant trend with primary cancers > hormone naïve lymph node metastases > hormone refractory autopsy metastases (P < 0.0032). These findings suggest that LINE1 hypomethylation occurs to some extent in primary prostate cancers compared with normal tissues but is more pronounced in metastatic prostate cancer.

A tiered gene-expression microarray and bisulfite sequencing–based approach to identify gene-specific CpG island hypomethylation changes in prostate cancer cells. Ever since the finding that DNA methyltransferase inhibitors can induce differentiation changes in cells (37), the notion that DNA methylation patterns are partly responsible for maintaining the differentiated phenotype has been very attractive. In support of this notion, recent studies have identified tissue differentially methylated genes that are selectively methylated in certain organs and cell types but not in others within individuals (13). More recently, a link to carcinogenesis was established by study findings that some genes that are normally methylated and silenced in specific tissues can become hypomethylated and overexpressed in cancers arising in that tissue (38–40).

Here, we used a tiered genome-wide gene expression microarray and bisulfite sequencing–based strategy to identify normally silenced and methylated genes that become expressed in prostate cancer cells due to CpG island hypomethylation. The central premise was that normally methylated genes that become hypomethylated in cancers may be identifiable as those genes that are absolutely unexpressed in normal cells but are expressed significantly in cancer cells. We therefore carried out gene expression microarray experiments on normal PrEC and stromal cells (PrEC, 4ST, 1S) and on six prostate cancer cell lines (LNCaP, C4-2B, LAPC-4, CWR22Rv1, PC-3, DU-145) using the Affymetrix U133A whole genome expression array platform. Next, we
Gene Expression
(Arbitrary units)

Extent of Methylation
by COMPARE assay
(Methylation Index)

Extent of Methylation
by Bisulfite Genomic Sequencing
(fraction of CpGs methylated)

1.0 0.5 0.0

Cancer Research
Cancer Res 2008; 68: (21). November 1, 2008 8960 www.aacrjournals.org
applied the recently developed Bar Code gene expression microarray analysis (22) to identify genes that were likely to be unexpressed in the normal cells and likely to be robustly expressed in at least one of the six cancer cell lines. Unlike most conventional microarray analytical methods, which are best suited to identifying relative patterns of expression (i.e., over or under expressed genes compared with a reference), Bar Code microarray analysis was designed to identify absolute patterns of expression (i.e., classification of a gene transcript as being absent or present). The Bar Code microarray analysis was recently shown to be a powerful classification tool (22).

Applying this Bar Code microarray analysis method with stringent thresholds for our set of normal and malignant prostate cell systems, we identified 168 probe sets that showed high confidence of expression in at least one prostate cancer cell line while showing high confidence of absence of expression in the normal PrEC and stromal cells (Fig. 3; Supplementary Table S2). Removal of redundant probe sets targeting the same gene and selection of genes with 5’ CpG islands gave a total of 120 unique genes with 5’ CpG islands that showed absolute expression in the cancer cells and lack of expression in the normal cells—a pattern consistent with cancer-specific gene CpG island hypomethylation. We noted that a surprisingly large fraction (18%) of these genes could be classified as sex chromosome–associated cancer-testis antigen (CTA) genes, so-named because of their expression in normal testis and many cancers but not in other normal tissues (41). Because many CTA genes have been previously found to become hypomethylated in cancers, we selected all of these genes (21 total), as well as another 10 randomly selected genes from a variety of functional categories, and successfully designed bisulfite sequencing assays for 22 genes, including 14 CTA genes and 8 other genes. Interestingly, multiple CTA genes often shared identical 5’ upstream CpG island sequences, both in direct and inverted repeat configurations. In some cases, the associated 5’ upstream CpG island was repeated more than 10 times across the X chromosome in the reference genome. All genes that were found to have methylated 5’ upstream CpG islands in the normal PrEC, 4ST, and/or WBC samples were sequenced in all of the six prostate cancer cell lines to see if they became hypomethylated at these CpG islands.

By this analysis, a total of 14 genes (CTAG1B, CTAG2, GAGE2, GAGE3, GAGE4, GAGE6, GAGE7, GAGE7B, MAGEA1, MAGEA3, MAGEA6, MAGEA12, PAGE1, TSPY1), containing seven unique 5’ upstream CpG island sequences, were found to be hypomethylated in at least one of the cancer cell lines compared with the normal cells (Fig. 4 and Supplementary Fig. S2). Expression of the majority of these hypomethylated CTA genes showed a strong inverse correlation with methylation levels (Fig. 4 and Supplementary Fig. S8), suggesting that hypomethylation of these genes is strongly associated with their overexpression in cancer cells. The only significant exceptions to this general trend were the GAGE genes, which all shared a single CpG island that was repeated more than 10 times on the X chromosome. Therefore, it is likely that this CpG island is either not involved in the regulation of these genes or plays a very complex role that is not easily deciphered. Surprisingly, all of these candidate hypomethylated and overexpressed genes belonged to the group of CTA genes, whereas none of the non-CTA genes tested showed any methylation in the normal WBC and PrEC cells and, therefore, by definition, could not be hypomethylated in the cancer cell lines because they were already unmethylated in the normal cells (Supplementary Fig. S3). It should be noted that these studies were not intended to identify all hypomethylated genes, because there are many other candidates from the original list that were not tested and use of less stringent thresholds would have identified many other candidates. Nonetheless, identification of this set of hypomethylated and overexpressed genes in the prostate cancer cell lines allowed subsequent analysis of the nature, extent, and timing of these representative gene CpG island hypomethylation changes in prostate cancer tissues.

We designed high-throughput COMPARE assays to assess the extent of CpG island hypomethylation at these candidate genes in prostate cancers. Results from these COMPARE assays were highly correlated with results from bisulfite sequencing in the prostate cell lines (Fig. 4 and Supplementary Fig. S8), suggesting that COMPARE was a good surrogate for bisulfite sequencing analysis, with the advantage that it was more amenable to high-throughput analysis. Application of these assays to prostate tissues revealed that, compared with the organ donor and tumor-adjacent benign PIA/PIN/normal prostate tissues, nearly all of the seven candidate CpG islands were hypomethylated in the primary but more so in the metastatic prostate cancers (Fig. 5A). Additionally, for six of the seven CpG islands tested, there was a statistically significant trend in the extent of methylation, with primary cancer > hormone naïve lymph node metastases > hormone refractory autopsy metastases. Data from microarray studies in the Oncomine database (42) showed that these hypomethylated genes were, in general, overexpressed at the mRNA level in prostate cancers compared with benign prostate and that the overexpression was often more pronounced in metastatic prostate cancers than in primary prostate cancers (Supplementary Fig. S4). Furthermore, immunohistochemical analysis of NY-ESO-1, encoded by the CTAG1B gene, and MAGEA1 in the prostate cancer cell lines showed that these proteins were only expressed in those cell lines that had significant hypomethylation at the 5’ CpG island of their corresponding gene (Fig. 5B and Supplementary Fig. S5). Immunohistochemical analysis of NY-ESO-1 in a prostate TMA containing 116 metastatic lesions, 26 primary cancers, and 19 benign prostate tissues showed that control testis tissues and a subset of metastases (13%) showed

Figure 4. Hypomethylation of gene-specific CpG islands is correlated with the expression of the corresponding genes. Right, extent of methylation at genes that were found to be hypomethylated in prostate cancer cell lines by bisulfite sequencing. For each sample, multiple alleles (number shown to the right of each sample) of each CpG island were cloned and sequenced after bisulfite modification and amplification. The frequency of methylation across all alleles for each CpG dinucleotide was color scaled from black (all alleles methylated) to light blue (none of the alleles methylated). A schematic of each CpG island, with position, relative to the transcriptional start site, of the start and end of the bisulfite sequencing amplicon and the position of each CpG (vertical hash marks) within each amplicon, is shown for each gene above the summary alleles for each specimen. These figures represent a summary of the complete data shown in Supplementary Fig. S2, which shows the pattern of methylation in each bisulfite sequenced clone individually. NA, not amplifiable. Middle, extent of methylation at the same genes as in the right panel as determined by the COMPARE assay. For each sample, MI is shown on the x axis, NA, the M.SssI-treated fraction for these specimens was not amplifiable, and therefore, these specimens were not informative. Data from COMPARE assays closely parallel data from bisulfite sequencing experiments. Left, extent of expression at the same genes as in the middle and right panels as determined by Affymetrix hgu133A gene expression microarrays. For each sample, the extent of methylation in normalized arbitrary expression units is shown on the x axis. When multiple genes contained identical 5’ CpG island sequences, expression levels of all associated genes are shown.
robust staining but that none of the primary cancers or benign prostate tissues showed any staining (Fig. 5C). Immunohistochemical analysis of MAGEA1 in prostate TMAs showed weak to moderate positive staining in ~3% to 5% of primary and metastatic prostate cancers and no staining in benign prostate tissues (data not shown). These findings are in general agreement with previous studies examining the expression pattern of these CTAs in the prostate (43). The significantly lower frequency of protein staining compared with the frequency of DNA hypomethylation and mRNA expression in the metastatic and primary
cancers may be an indication that mRNA expression and/or CpG island hypomethylation is not sufficient for protein expression for these genes in prostate cancer tissues. Alternatively, it is possible that the low degree of staining in the prostate cancer tissues compared with the cell lines may reflect a limitation in the ability to detect these proteins by immunohistochemistry in prostate tissue sections, in which antigen preservation and presentation may have been less robust than in freshly prepared and fixed cell line specimens. Evaluation of commercially available antibodies for proteins encoded by the other hypomethylated genes, including MAGEA3, MAGEA6, MAGEA12, and TSPY1, showed that immunohistochemical analyses with these antibodies were highly sensitive to tissue preparation and fixation conditions, often showed high nonspecific signal and low dynamic range, and were, therefore, not well suited for immunohistochemical analysis of our TMAs (data not shown).

These results suggest that, whereas some gene specific CpG island hypomethylation does occur in primary prostate cancers, these DNA hypomethylation changes, like LINE1 promoter hypomethylation and overall 5mC content reductions, are most prominent in the metastatic prostate cancers.

DNA hypomethylation changes are more heterogeneous across different anatomically distinct metastatic deposits within the same patients than characteristic CpG island hypermethylation changes. Yegnasubramanian and colleagues previously showed that nearly all the anatomically distinct metastatic deposits from a given subject showed very similar CpG island hypermethylation patterns, suggesting that these changes were present in the primary cancer and were closely maintained during and after metastatic dissemination (18). Here, we examined whether DNA hypomethylation changes were also closely maintained during the process of metastatic dissemination by analyzing the same metastatic deposits that were examined in the previous study of CpG island hypermethylation. Compared with the CpG island hypermethylation changes, DNA hypomethylation changes were much more heterogeneous across different anatomically distinct metastatic deposits from the same subject (Fig. 6A). To confirm this visual observation, we used linear regression models to fit the CpG island hypermethylation and CTA gene hypomethylation data using patients, DNA methylation markers, and patient-marker interactions as independent variables. In addition, we fit the LINE1 promoter methylation and 5mC content data using patients as the independent variable. All of these models adequately fit the data, accounting for ~71% and ~91% of the variance in the CTA gene hypomethylation and CpG island hypermethylation data, respectively, and for ~81% and ~67% of the LINE1 promoter hypomethylation and 5mC content data, respectively. Examination of the residual variances, which represent the variance across sites within each patient or patient-marker combination, revealed that the variance in the hypomethylation data across sites within each patient was significantly higher than the variance in the CpG island hypermethylation data by the F test (5mC content, 2.1-fold higher variance, P < 10^{-5}; LINE1 promoter methylation, 1.7-fold higher variance, P < 10^{-3}; CTA gene hypomethylation, 2.1-fold higher variance, P < 10^{-10}). Additionally, the pattern of 5mC immunohistochemical staining in these metastatic deposits was very heterogeneous across metastatic deposits from the same subject and even seemed to be quite focal and heterogeneous within a single metastatic deposit in some samples (Fig. 6B). Applying the same type of regression analysis, the variance in the 5mC immunohistochemical data across sites within patients was >3.5-fold higher than that in the CpG island hypermethylation data (P < 10^{-10}). Finally, consistent with the pattern of hypomethylation in the metastatic lesions, immunohistochemical staining of these metastatic deposits for NY-ESO-1 showed that its expression was also heterogeneous across different metastatic deposits within the same patient and often showed focal heterogeneous expression within a single metastatic deposit (Fig. 6C). This CTA gene expression heterogeneity is consistent with that observed previously even in primary cancers (41) and suggests that NY-ESO-1 expressing cells from the primary cancers were unlikely to have been clonally selected for during the process of metastatic dissemination because metastases also show considerable heterogeneity in their expression of NY-ESO-1. Additionally, DNA hypomethylation changes, like the CpG island hypermethylation changes, did not show any site-specific patterns but were rather heterogeneous across patients within each site category (Supplementary Fig. S6). These data suggest that DNA hypomethylation changes occur most prominently during or after metastatic dissemination and contribute to tumor heterogeneity in metastatic prostate cancer.

A model of DNA methylation changes in prostate cancer initiation and progression and its implications for disease pathogenesis and treatment. Several studies have previously shown that CpG island hypermethylation changes occur nearly universally in prostate cancer, often arise early in prostate cancer initiation and progression, and, once they arise, are closely maintained throughout disease progression and metastatic dissemination (18, 44). In contrast, here, we show that DNA hypomethylation changes occur late in prostate cancer progression, occurring most frequently in metastatic lesions, and are often heterogeneous across different metastatic deposits within the same subjects. Because others have reported age-related changes in DNA methylation (e.g., refs. 45, 46), we examined whether the hypomethylation changes between groups observed in this study were independent of age. These DNA hypomethylation changes were not simply a function of increasing age in the men with advanced prostate cancer because nearly all statistically significant

Figure 5. Gene-specific CpG island hypomethylation occurs in primary prostate cancer but is more pronounced in metastatic prostate cancer. A, box-and-whisker plots showing MI as determined by COMPARE assay at each of the CpG islands found to be hypomethylated in the prostate cancer cell lines. The results of t test for all statistically significant differences in the mean 5mC content between all pairwise combinations within each gene are shown. * P < 0.05; ** P < 0.005; *** P < 0.0005. Statistically significant trends in CTA gene methylation with primary > hormone naive lymph node metastases > hormone refractory autopsy metastases are indicated. 1, P < 0.05; 2, P < 0.005; 3, P < 0.0005. Differences remained statistically significant after age adjusting. The number of specimens analyzed in each prostate tissue category is indicated in the legend. 4, for the autopsy metastases, 76 metastatic deposits from 28 subjects (one to seven metastatic deposits from each subject) were analyzed. B, immunohistochemistry with anti–NY-ESO-1 antibodies in prostate cancer cell lines. NY-ESO-1 protein (encoded by the CTAG1B gene) staining was seen in cell lines that had a high degree of hypomethylation at this gene CpG island (CWR22Rv1 and LAPC-4) but not in the other prostate cancer cell lines. C, immunohistochemistry with anti–NY-ESO-1 antibodies in prostate tissues and control normal tissues. No NY-ESO-1 protein staining was observed in any of the benign prostate and primary prostate tissues tested. Robust staining was seen in the testis and in 13% of metastatic prostate cancers. Representative positive staining sections of metastatic prostate tissues and normal testis. Boxed regions are shown in higher magnification. Original optical magnifications are indicated.
differences in $^{5mC}$ content, LINE1 hypomethylation, and CTA gene hypomethylation between different tissue categories remained significant even after adjusting for age in multivariate regression models.

Interestingly, the tumor-adjacent benign prostate tissues containing areas of PIN and/or PIA often contained low levels of CpG island hypermethylation changes (18, 23) but did not have any apparent reductions in $^{5mC}$ content or hypomethylation of LINE1 and CTA gene promoters. Additionally, whereas the CpG island hypermethylation changes examined previously occurred nearly universally in primary prostate cancer, only a fraction of these primary cancers showed decreases in LINE1 and CTA gene promoter methylation. These findings suggest that these hypomethylation changes occur later in prostate cancer progression than the CpG island hypermethylation changes and may indicate that DNA hypomethylation changes are not important in prostate cancer initiation but may be more important in the formation/progression of lethal metastatic prostate cancer. This is in stark contrast to findings in other cancers where DNA hypomethylation reportedly occurs early in carcinogenesis and contributes to initiation of the neoplastic phenotype, occurring frequently and prominently in precancerous lesions of organs, such as lung and colon (9, 15).

Instead, the findings here suggest that most of the DNA hypomethylation changes examined in this study were highly
correlated with lethal metastatic prostate cancer. Interestingly, whereas significant $^{3}\text{H}\text{C}$ content reductions occurred only at the stage of metastatic prostate cancer, LINE1 promoter hypomethylation and CTA gene hypomethylation appeared first in primary prostate cancers and became more pronounced in metastatic prostate cancers. Whereas these observations may suggest that LINE1 hypomethylation or CTA gene hypomethylation in primary cancers may lead to a selective advantage for the development of metastatic disease, two major findings in this study provide some evidence against this hypothesis. First, whereas CpG island hypermethylation changes were uniformly maintained across multiple metastases from the same individual, indicating that these arose before metastatic dissemination, DNA hypomethylation changes were much more variable across different metastases from the same individual, suggesting that these changes commonly arose after metastatic dissemination. Second, whereas various CpG island hypermethylation changes were highly associated with pathologic stage, grade, and risk of recurrence (18), none of the DNA hypomethylation changes in primary cancers observed here were associated with pathologic grade, stage, or recurrent disease (data not shown). These findings suggest that, although the LINE1 and CTA promoter hypomethylation sometimes occur in the primary cancers, these hypomethylation changes do not confer an increased risk for aggressive disease and are not necessarily selected for and maintained during metastatic dissemination. It is also interesting to note that the primary cancers had a statistically significant increase in the global $^{3}\text{H}\text{C}$ content compared with the organ donor normal prostate specimens, despite having decreases in LINE1 promoter and CTA gene methylation. This suggests that hypomethylation changes occurring in LINE1 promoter regions and CTA gene 5’ upstream regions in these primary cancers are offset and

Figure 6  Continued. C, representative images of NY-ESO-1 immunohistochemistry in multiple metastases each from different subjects. The degree of NY-ESO-1 staining was highly heterogeneous in different metastatic deposits within each subject. Multiple deposits from two representative subjects. The original optical magnification is indicated. D, a model of DNA methylation changes in prostate cancer initiation and disease progression.
sometimes overbalanced by hypermethylation changes in other portions of the genome, likely at numerous CpG islands.

These observations imply that the DNA hypomethylation changes occur most consistently during or after metastatic dissemination, at which point they may even contribute to the maintenance and propagation of the metastatic lesions. Several recent reports have shown that increased hypomethylation is highly correlated with genetic instability in human malignancies (9–11). The increased hypomethylation observed in the metastatic prostate cancer lesions may contribute to the high number and prevalence of chromosomal aberrations previously reported in metastatic prostate cancers compared with primary cancers (47). One recent report synthesizing data from 41 studies analyzing 872 cancer specimens showed that advanced prostate cancer specimens show a significantly higher frequency of chromosomal abnormalities than primary prostate cancer specimens (48). Consistent with this hypothesis is the finding that mice with hypomorphic DNMT enzymes are prone to genetic instability and lymphomagenesis (49).

The previous findings regarding CpG island hypermethylation alterations in prostate cancer highlighted a fundamental paradox: there was enough dysregulation of the DNA methylation machinery to produce the initial CpG island hypermethylation changes, and yet, there was enough fidelity in the DNA methylation machinery so that a specific set of CpG island hypermethylation changes could be closely maintained throughout disease progression even through metastasis. These observations were consistent with the notion that there is an early dysregulation of the de novo DNA methylation machinery but that the maintenance DNA methylation machinery remains intact, allowing for maintenance of these changes throughout disease progression. Yet, the data obtained here suggests that this cannot be completely accurate. Indeed, there seems to be a progressive dysregulation of maintenance methylation, manifesting most prominently late in disease and causing reductions in overall 5mC content, in LINE1 methylation and in gene-specific CpG island methylation. These data suggest a revised model for dysregulation of DNA methylation in prostate cancer: increased de novo methylation early in prostate carcinogenesis to produce coordinate CpG island hypermethylation changes, with progressive loss of maintenance methylation in the midst of selective pressures to maintain hypermethylation of CpG islands in the 5′ upstream region of several key genes (Fig. 6D). Such a model could explain the early and frequent appearance of several CpG island hypermethylation changes in prostate cancer, the ability to maintain these changes at key CpG islands during disease progression, and the loss of overall 5mC content, LINE1 methylation, and gene-specific CpG island methylation occurring late in prostate cancer progression.

Tumor cell heterogeneity, particularly in the expression and function of therapeutic targets, is a common feature in advanced cancers that poses major challenges in the development of effective treatments. Here, we show that DNA hypomethylation is a significant source of tumor heterogeneity in metastatic prostate cancer and may contribute to the development of therapeutic resistance. In particular, we show that gene-specific DNA hypomethylation changes can cause heterogeneous overexpression of a series of cancer-testis antigen genes, many of which are currently being evaluated as targets for immunotherapy (41, 50). Data from recent clinical trials have shown regression of tumors when patients were treated with immunotherapies targeted to these CTA antigens (51, 52). However, because there is significant heterogeneity in the DNA hypomethylation and CTA gene expression patterns from tumor to tumor in metastatic prostate cancer, vaccine strategies directed against a single CTA target may be of limited utility, unless these strategies can also produce more promiscuous immune responses to eradicate nonexpressing tumor cells, as recently described (53). Based on the heterogeneity of DNA hypomethylation and associated gene overexpression shown in this study, further studies to test the potential of vaccine strategies directed against multiple CTA gene products in combination or use of vaccines containing multiple heterogeneous whole tumor cell models may yield new therapies for advanced prostate cancer (54).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 11/2/2007; revised 7/25/2008; accepted 8/12/2008.

Grant support: NIH/National Cancer Institute grants CA070196, CA113374, CA98286, and CA122344, Department of Defense PCRD grant PC073533, and the Prostate Cancer Foundation.

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.

We thank Jeanne Kowalski and Amanda Blackford for providing advice with regards to analysis of the gene expression microarray data, Kristen Leckssell for assistance in scanning TMAs, and the Johns Hopkins TMA facility for preparation and processing of TMAs.

References


33. Nair I, Pascale E, Furano AV. The left end of rat L1 (L1Rn, long interspersed repeated) DNA which is a CpG island can function as a promoter. Nucleic Acids Res 1988;16:9233–51.


DNA Hypomethylation in Prostate Cancer Progression