DNA methylation profiles delineate etiologic heterogeneity and clinically important subgroups of bladder cancer

C.S. Wilhelm-Beinartzi1, D.C. Koestler1, E.A. Houseman1, B.C. Christensen1,2, John K. Wiencke3, A.R. Schmed4, M.R. Karagas5, K.T. Kelsey1 and C.J. Marsit1,2,*

1Department of Community Health, Center for Environmental Health and Technology and 2Department of Pathology and Laboratory Medicine, Brown University, Providence, RI 02912, USA, 3Department of Neurological Surgery, University of California-San Francisco, San Francisco, CA 94143, USA and 4Department of Pathology and 5Department of Community and Family Medicine, Dartmouth Medical School, Hanover, NH 03756, USA

*To whom correspondence should be addressed. Department of Pathology and Laboratory Medicine, Brown University, Box G-E537, Providence, RI 02912, USA. Tel.: +1 401 863 6508; Fax: +1 401 863 9008; Email: carmen_marsit@brown.edu

DNA methylation profiles can be used to define molecular cancer subtypes that may better inform disease etiology and clinical decision-making. This investigation aimed to create DNA methylation profiles of bladder cancer based on CpG methylation from almost 800 cancer-related genes and to then examine the relationship of those profiles with exposures related to risk and clinical characteristics. DNA, derived from formalin-fixed paraffin-embedded tumor samples obtained from incident cases involved in a population-based case-control study of bladder cancer in New Hampshire, was used for methylation profiling on the IInulina GoldenGate Methylation Bead Array. Unsupervised clustering of these loci with the greatest change in methylation between tumor and non-diseased tissue was performed to define molecular subgroups of disease, and univariate tests of association followed by multinomial logistic regression was used to examine the association between these classes, bladder cancer risk factors and clinical phenotypes. Membership in the two most methylated classes was significantly associated with invasive disease \( P < 0.001 \) for both class 3 and 4. Male gender \( (P = 0.04) \) and age \( > 70 \) years \( (P = 0.05) \) was associated with membership in one of the most methylated classes. Finally, average water arsenic levels in the highest percentile predicted membership in an intermediate methylated class of tumors \( P = 0.02 \) for both classes. Exposures and demographic associated with increased risk of bladder cancer specifically associate with particular subgroups of tumors defined by DNA methylation profiling and these subgroups may define more agressive disease.

Introduction

Identification of molecularly defined subgroups of tumors holds the promise of personalized treatment strategies (1). For example, examination of RNA expression in a panel of genes, in breast cancer, is now used clinically to provide more individualized, targeted and less toxic forms of therapy (2). In addition, for understanding cancer etiology, the examination of molecular profiles of tumors has demonstrated considerable utility in delineating carcinogen exposure-associated differences in individual tumors (3,4). Bladder cancer is the ninth most incident form of cancer in the USA with over 70 000 new cases diagnosed in 2009 (5). Seventy percent of bladder cancers are non-invasive and highly treatable, although these are more probably to recur (6). Thirty percent of bladder cancers are invasive at presentation, spreading into and through the muscular layers of the bladder and causing high rates of death from metastasis (6,7). This cancer is three to four times more common in men with tobacco smoking being the main risk factor for this disease. Other risk factors include occupational exposures, arsenic ingestion, chlorination by-products and possibly hair dye use and dietary factors (8–10).

Epigenetics is an evolving research area with potential utility for apportioning etiologic fractions as well as for designing future personalized therapies. Epigenetics involves heritable stable changes to gene expression, which are potentially reversible. These changes include DNA hypermethylation leading to gene silencing as well as DNA hypomethylation, leading to oncogene activation and genomic instability (11,12). Alterations in the DNA methylation pattern of the promoter region of cancer-related genes have been associated with risk factors, clinical presentation and outcomes of bladder cancer (13). These risk factors include smoking, arsenic, age and gender, all of which have been associated with an increased prevalence of individual gene alterations or coordinated epigenetic alteration of a small panel of genes in bladder tumors (14–18). Expanding on this concept, previous work in breast, colorectal and head and neck cancer has shown that profiles of the gene promoter methylation may define type of disease and have been associated with the etiologic and clinicopathological features of those diseases (19–21).

In this study, we sought to utilize the DNA methylation profiles of bladder cancers based on the CpG methylation of almost 1500 CpG loci associated with >800 cancer-related genes to identify molecular subgroups of the disease. We then examined the association of those subgroups with risk factors of bladder cancer in order to gain an improved understanding of the etiology of this disease. This approach may help better target prevention efforts and aid in identifying novel subtypes of bladder cancer of therapeutic interest.

Materials and methods

Subjects

A description of the study design appears in earlier reports (22,23). Briefly, bladder cancer cases were drawn from subjects enrolled in two stages of a non-consecutive population-based case-control study of bladder cancer in New Hampshire, conducted from 1994 to 1998 and from 2001 to 2004. Cases of incident bladder cancer were identified from the state cancer registry and a standardized histopathologic review was conducted by a single study pathologist (A.R.S.) to verify the diagnosis and histopathology of the cases. Formalin-fixed paraffin-embedded tumor tissue was obtained from a subset of the cases in the overall study. In addition, non-diseased bladder epithelium \( (n = 5) \) was obtained from individuals without cancer through the National Disease Research Interchange. All of these samples came from men, with ages of 22, 68, 72, 75 and 84 years, with four of five being smokers. For the analyses presented here, the case group was restricted to Caucasian transitional cell carcinomas having smoking status data and promoter methylation data and excluded cases that were diagnosed as carcinoma in situ due to small numbers; this included a total of 310 cases \( (n = 53 \) from series 1 of 459 cases and \( n = 257 \) from series 2 of 398 cases), whose characteristics are presented in Table I. Ninety-five percent of cases and controls in this study were Caucasian, giving us limited power to detect differences in bladder cancer risk factors or prognosis in other racial/ethnic groups; therefore, we restricted our analyses to only Caucasians with race/ethnicity being obtained through self-report. For efficiency purposes for phase 1, the same control group used in a study of non-melanoma skin cancer conducted from 1 July 1993 to 30 June 1995 was used (24). Additional controls were selected afterward up to 2002 frequency matched to cases by age \( (25–34, 35–44, 45–54, 55–64, 65–69 \) and \( 70+) \) and gender and randomly assigned a reference date comparable with the cases' diagnosis date as described previously (22). In both series, controls \(<65\) years of age were selected from records obtained from the New Hampshire Department of Transportation and controls \(>65\) years of age were chosen from records obtained from the Health Care Financing Administration’s Medicare Program. Approximately 70% controls eligible for this study agreed to participate and the same methods were used to collect controls in all phases of the study. For the analyses presented here, the control group was restricted to 1546 Caucasian controls \((n = 637 \) from series 1 and \( n = 909 \) from series 2) with complete smoking data, whose characteristics are presented in Table I. No
Table I. Selected characteristics of bladder cancer cases stratified by RPMMs class and controls with smoking status data

<table>
<thead>
<tr>
<th>Covariates</th>
<th>All cases (N = 1856)</th>
<th>Controls (N = 310), n (%)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoke status (N = 1856)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>53 (17.1)</td>
<td>537 (34.73)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Former</td>
<td>161 (51.94)</td>
<td>751 (48.58)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>96 (30.97)</td>
<td>258 (16.69)</td>
<td></td>
</tr>
<tr>
<td>Gender (N = 1856)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>75 (24.19)</td>
<td>561 (36.29)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Male</td>
<td>235 (75.81)</td>
<td>985 (63.71)</td>
<td></td>
</tr>
<tr>
<td>Age (N = 1856)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 years</td>
<td>27 (8.71)</td>
<td>233 (15.07)</td>
<td>&lt;0.0004</td>
</tr>
<tr>
<td>50–59 years</td>
<td>60 (19.35)</td>
<td>296 (19.15)</td>
<td></td>
</tr>
<tr>
<td>60–69 years</td>
<td>100 (32.26)</td>
<td>569 (36.8)</td>
<td></td>
</tr>
<tr>
<td>70–79 years</td>
<td>123 (39.68)</td>
<td>448 (28.98)</td>
<td></td>
</tr>
<tr>
<td>Water Arsenic (N = 1483)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002 to &lt;0.104 μg/l</td>
<td>56 (18.79)</td>
<td>297 (25.06)</td>
<td>0.02</td>
</tr>
<tr>
<td>0.104–160.50 μg/l</td>
<td>242 (81.21)</td>
<td>888 (74.94)</td>
<td></td>
</tr>
</tbody>
</table>

Results

The characteristics of controls and cases used in analyses are shown in Table I. We initially examined the difference in the profiles of methylation between the non-diseased bladder epithelium and bladder tumors (supplementary Figure 1A is available at Carcinogenesis Online). An RPMM utilizing all autosomal loci comparing tumor vs. non-diseased tissue found that all of the non-diseased tissue clustered into a single class, and thus the profiles of DNA methylation are significantly different between tumor and non-diseased tissue (supplementary Figure 1B is available at Carcinogenesis Online; P < 0.00001).

Figure 1 depicts the DNA methylation data for four distinct classes of bladder tumors resulting from RPMM. We found 267 loci that had an absolute median delta-beta value that met the >0.2 threshold so we then fit a beta distributed RPMM to the tumor samples including only these loci. The automated RPMM solution resulted in nine classes; however, due to the small number of participants in some classes, we combined the smallest classes with their RPMM siblings, effectively pruning the RPMM dendrogram further than the automated solution to four final RPMM classes, resulting in four classes. The intensity of methylation is shown in the heatmap with yellow indicating unmethylated and blue indicating fully methylated. Overall, class-specific means of the delta beta across all loci was lowest in class 1 (mean average beta = 0.25), followed by class 2 (mean average beta = 0.35), then class 3 (mean average beta = 0.44) and class 4 was the most highly methylated (mean average beta = 0.48). The proportion of non-invasive and invasive cases by RPMM class is also shown in Figure 1. The association between having invasive bladder cancer and RPMM class was found to be significant in a permutation chi-square test at the P < 0.05 level.

Table II presents the odds of invasive bladder cancer by methylation class, adjusted for gender, age and smoking status using unconditional logistic regression. Compared with class 1, a significant 3.93-fold increased risk of being invasive [95% confidence interval (CI):...
1.96–7.89) was observed among subjects in class 3, and a significant 4.89-fold increased risk of being invasive (95% CI: 2.15–11.09) was found in subjects in class 4.

To look at the potential associations between bladder cancer risk factors and specific methylation class, we used controls as our referent group and ran a multinomial logistic regression on all four classes against controls (Table III) and as a stratified analysis using unconditional logistic regression on each class versus controls (supplementary Table 1 is available at Carcinogenesis Online). We only included, in our final model, covariates found to be significant at the \( P < 0.05 \) level in previously run univariate analyses as well as the matching factors of age and gender. Current smokers had similar, significant odds of membership in any methylation class compared with controls (Table III; \( P < 0.05 \)), whereas former-smokers were significantly associated with membership in class 2 and 3. Being male significantly odds of membership in any methylation class compared with class 3 [odds ratio (OR) 1.75, 95% CI: 1.04–2.95] along with being >70 years of age (OR 2.28, 95% CI: 1.02–5.08). Finally, cases in the highest quartile of average water arsenic levels significantly predicted membership in classes 2 and 3 with an OR of 2.02 (95% CI: 1.12–3.63) and 1.98 (95% CI: 1.09–3.60), respectively. Further, the stratified analysis presented in supplementary Table 1 (available at Carcinogenesis Online) is consistent with the results of the multinomial regression.

### Table II. Risk of invasiveness for each RPMMs class adjusting for smoke status, age, gender and average water arsenic levels

<table>
<thead>
<tr>
<th>RPMM Class ( (N = 310 \text{ cases})^a )</th>
<th>Non-invasive ( (n = 90), \text{ n (%)} )</th>
<th>Invasive ( (n = 211), \text{ n (%)} )</th>
<th>OR for invasive disease (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 ( (n = 84) )</td>
<td>68 (81.0)</td>
<td>16 (19.0)</td>
<td>Ref</td>
</tr>
<tr>
<td>Class 2 ( (n = 91) )</td>
<td>76 (83.5)</td>
<td>15 (16.5)</td>
<td>0.97 (0.44–2.10)</td>
</tr>
<tr>
<td>Class 3 ( (n = 82) )</td>
<td>44 (53.7)</td>
<td>38 (46.3)</td>
<td>3.93 (1.96–7.89)</td>
</tr>
<tr>
<td>Class 4 ( (n = 53) )^b</td>
<td>23 (52.3)</td>
<td>21 (47.73)</td>
<td>4.89 (2.15–11.09)</td>
</tr>
</tbody>
</table>

Italicized values represent statistically significant associations.

^aThe model is adjusted for smoke status, age, gender and average water arsenic levels.

^bClass 4 is missing nine stage or grade values.

Discussion

This study utilized methylation profiles to define subtypes of bladder cancer and associated these subtypes with clinical disease presentation and carcinogen exposure histories. As expected, our initial analyses demonstrate that the profiles identified in tumors are significantly distinct from those identified in non-diseased bladder epithelium. Among tumors, we demonstrated that the mean methylation level differs among methylation profile classes, suggesting that there are distinct phenotypes associated with the methylation profiles, and that membership in the most methylated classes is associated with ORs for invasive bladder cancer of \( \geq 4 \). This is consistent with our initial analyses using a highly selected reduced number of loci in a smaller series of tumors, which demonstrated that a greater propensity for DNA methylation was associated with more aggressive forms of bladder cancer (15). Our previous work also suggested that the propensity identified by a small number of genes may in fact have been exemplifying a more widespread process of epigenetic dysregulation across the genome (15).

Again, consistent with this previous report and previously published work (14–17,34,35), we also have demonstrated associations between male gender, age and former smoking status with specific subgroups of bladder tumors defined by methylation profile. Compared with non-smokers and controls, current-smokers demonstrated relatively similar odds of membership in all methylation-based subgroups of bladder cancer. This suggests that the specificity of class membership is based on additional exposures, beyond current smoking. For example, high water arsenic levels were associated with cases that had a class 3 methylation profile, suggesting that arsenic exposure has a distinct phenotype represented by a highly specific epigenetic profile. Arsenic exposure has been associated with epigenetic effects in animal models (36–39) and we have demonstrated that specific gene methylation events are associated with arsenic exposure in bladder cancer (14). This class was also almost four times more probably to be an invasive tumor compared with class 1, consistent with our findings that arsenic exposure is associated with more aggressive disease and poorer patient survival (28). Although there is controversy over the levels at which arsenic exposure is carcinogenic in humans, our data suggest that levels found commonly in the USA (40,41) give
DNA methylation profiles of bladder cancer

Table III. Nominal logistic regression of association of covariates and risk of being in a specific RPPM class using controls as the referent group

<table>
<thead>
<tr>
<th>Class 4, OR (95% CI)</th>
<th>Class 3, OR (95% CI)</th>
<th>Class 2, OR (95% CI)</th>
<th>Class 1, OR (95% CI)</th>
<th>Controls, N = 1546</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 cases, n (%)</td>
<td>5 cases, n (%)</td>
<td>8 cases, n (%)</td>
<td>11 cases, n (%)</td>
<td></td>
</tr>
<tr>
<td>0.30 (0.16–0.55)</td>
<td>0.36 (0.18–0.72)</td>
<td>0.66 (0.33–1.34)</td>
<td>0.97 (0.48–1.98)</td>
<td></td>
</tr>
</tbody>
</table>

Note: controlling for factors significantly associated with disease.

In summary, this study demonstrates that profiles of DNA methylation can be used to distinguish phenotypically and clinically important subgroups of bladder cancer. Smoking history as well as arsenic exposure, age and gender are not only risks of bladder cancer in general but also predispose individuals to specific molecular subtypes of disease. The novelty of these results lies in the use of array-based methodologies to examine CpG methylation of a large number of CpG loci instead of examining only specific promoter regions of certain genes, thereby allowing for a more comprehensive understanding of the epigenetic landscape of bladder tumors. These findings indicate that the methylation profiles of CpG loci can be used as a potential diagnostic marker of bladder cancer and can help further identify novel molecular subtypes of bladder cancer. Future work should examine if these subtypes can be used to create more individualized, targeted regimens of therapy for bladder cancer and aid in the prognosis of this disease.
Conflict of Interest Statement: None declared.

References


Received May 19, 2010; revised August 12, 2010; accepted August 14, 2010

1976