Identification of Methylated Genes Associated with Aggressive Bladder Cancer

Carmen J. Marsit[1,2]*, E. Andres Houseman[2], Brock C. Christensen[1,2], Luc Gagne[1], Margaret R. Wrensch[3], Heather H. Nelson[4], Joseph Wiemels[3], Shichun Zheng[3], John K. Wiencke[3], Angeline S. Andrew[5], Alan R. Schned[6], Margaret R. Karagas[5], Karl T. Kelsey[1,2]

1 Department of Pathology and Laboratory Medicine, Brown University, Providence, Rhode Island, United States of America, 2 Department of Community Health, Center for Environmental Health and Technology, Brown University, Providence, Rhode Island, United States of America, 3 Department of Neurological Surgery, University of California San Francisco, San Francisco, California, United States of America, 4 Division of Epidemiology and Community Health, Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota, United States of America, 5 Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Dartmouth Medical School, Lebanon, New Hampshire, United States of America, 6 Department of Pathology, Dartmouth Medical School, Lebanon, New Hampshire, United States of America

Abstract

Approximately 500,000 individuals diagnosed with bladder cancer in the U.S. require routine cystoscopic follow-up to monitor for disease recurrences or progression, resulting in over $2 billion in annual expenditures. Identification of new diagnostic and monitoring strategies are clearly needed, and markers related to DNA methylation alterations hold great promise due to their stability, objective measurement, and known associations with the disease and with its clinical features. To identify novel epigenetic markers of aggressive bladder cancer, we utilized a high-throughput DNA methylation bead-array in two distinct population-based series of incident bladder cancer (n = 73 and n = 264, respectively). We then validated the association between methylation of these candidate loci with tumor grade in a third population (n = 245) through bisulfite pyrosequencing of candidate loci. Array based analyses identified 5 loci for further confirmation with bisulfite pyrosequencing. We identified and confirmed that increased promoter methylation of *HOXB2* is significantly and independently associated with invasive bladder cancer and methylation of *HOXB2*, *KRT13* and *FRZB* together significantly predict high-grade non-invasive disease. Methylation of these genes may be useful as clinical markers of the disease and may point to genes and pathways worthy of additional examination as novel targets for therapeutic treatment.

Introduction

In the United States in 2009, an estimated 71,000 cancers of the urinary bladder were diagnosed and greater than 14,000 deaths were attributed to this disease [1]. The vast majority of deaths occur in patients with incident high stage, high grade, invasive tumors that infiltrate the muscular layers of the bladder. Low grade, non-invasive disease, on the other hand, can be successfully treated, though this success comes at great economic burden to the healthcare system. Approximately 500,000 patients require monitoring in the U.S. leading to estimated diagnosis to death per patient costs ranging from $96,000 to $187,000, thereby resulting in $2.2 billion in annual expenditures, making bladder cancer the most expensive of all cancers [2,3]. Thus, cost-effective prognostic strategies for evaluating incident and recurrent disease would be of significant clinical utility.

Epigenetic control of DNA expression is well known to drive fetal developmental differentiation. In a parallel fashion, in concert with genetic events (mutation, deletion and gene amplification) it is thought that epigenetic alterations may precipitate important pathological features of malignant degeneration [4]. Bladder cancer, with its divergent clinical (and pathological) phenotypes, presents a tumor model that may arise by inactivation of loci that independently control the propensity for invasion and, hence, dictate stage and grade, and that this inactivation may occur through a variety of epigenetic processes including microRNA alterations [5], alterations to chromatin [6,7], and alterations to DNA methylation [8]. In this case, there is potential for the use of epigenetic alterations and particularly DNA CpG methylation as biomarkers for bladder cancers, as well as, potentially, for a variety of other human cancers [8,9,10,11,12,13,14]. Microarray-based approaches also have attempted to identify novel genes associated with invasive disease but with limited sample sizes due to the array strategy employed [15]. Recent developments in array approaches now allow for application of these technologies to population-based epidemiologic studies of cancer utilizing large numbers of samples [16,17,18]. There are numerous advantages to utilizing a population-based approach, including reduction in bias, greater generalizability of the results, and access to samples spanning all stages and grades of the tumor. Therefore, we have utilized this array-based approach to identify clinically and biologically informative patterns and novel gene targets of DNA CpG

methylation in a population-based series of bladder transitional cell carcinoma.

Results

Identification of candidate loci

We utilized a 2-step approach to the identification and validation of loci that were associated with the development of invasive bladder cancer (Figure S1). In step 1, two independent series of tumors (Table 1) were analyzed by DNA methylation array to identify potential candidate loci having differential methylation, and in step 2, these candidates were confirmed in an additional series of tumors not profiled on the array.

Overall, there was a general increase in methylation in invasive tumors (Figure 1A), suggesting that this approach has utility for demarcation of effective biomarkers. Generalized linear models comparing methylation levels of invasive versus non-invasive tumors at each locus confirmed this visual impression. In series 1, 445 CpG loci had significantly increased methylation in invasive compared to non-invasive tumors ($Q<0.05$) while only 68 loci had significantly decreased methylation. Similarly, in series 2, 606 loci had significantly increased methylation and only 41 had significantly decreased methylation (Table S1). Modeling each series, independently, with recursively partitioned mixture models yielded four methylation profiles (Fig 1B). Notably, significantly more invasive tumors were in methylation class 4 ($P<0.00001$, permutation chi-square).

To identify the loci most robustly associated with invasive disease, we utilized 3 distinct, commonly applied statistical approaches to identify those gene loci (from each of the three statistical approaches and across the 2 series) that overlapped in differentiating invasive and non-invasive disease. Five loci (FRZB_E186, HOXB2_P99, KRT13_P676, RIPK1_P868, STAT5A_P704) were found to be associated with invasive disease in all three statistical approaches.

For four of these loci (FRZB, STAT5A, KRT13, and HOXB2), bisulfite pyrosequencing assays were able to be designed, and were used to confirm the array findings in a subset of bladder tumors examined on the array. A pyrosequencing assay for RIPK1 could not be successfully designed. For all CpG sites examined, as well as the mean across sites for FRZB, KRT13, and HOXB2, we observed significantly greater methylation in invasive compared to non-invasive tumors, consistent with the array results (Figure S2A,

S2C, S2D). For STAT5A, we confirmed the significantly greater extent of methylation at the specific CpG site measured by the array, but did not observe this association in the neighboring CpG sites (Fig S2B), and so this loci was not further investigated.

Treatment of a Bladder Cancer Cell Line with Methylation Inhibitor Leads to Re-expression of HOXB2

Pyrosequencing was performed to determine the methylation status of HOXB2, FRZB, and KRT13 in bladder cancer cell lines HTB-9 and UM-UC3. HTB-9 had a HOXB2 methylation extent of 68.9, similar to what had been observed amongst primary bladder tumor samples, and thus was chosen for further examination. Treatment of HTB-9 cells with 1 or 2 μM 5-aza-2′-deoxycytidine led to an increase of greater than 100 fold expression of HOXB2 compared to mock-treated cells (Figure 2).

HOXB2 Methylation is independently associated with invasive bladder cancer

Pyrosequencing of the promoter regions of FRZB, KRT13, and HOXB2 was performed in an independent series of 263 bladder tumor patients not examined on the array as well as on 4 non-diseased tumor samples obtained from the NDRI. The extent of methylation at the pyrosequenced regions in non-diseased bladder epithelium, non-invasive tumors to invasive tumors is depicted on Figure 3, and reveals significant differences in methylation across the 3 categories for each of the genes (HOXB2 $P<0.00001$, KRT13 $P<0.05$, FRZB $P<0.003$, Kruskal Wallis Test). Methylation of HOXB2 and KRT13 are each significantly greater in invasive compared to non-invasive tumors (protected Wilcoxon rank sums test, $P<0.00001$ and $P<0.02$ respectively). TP53 immunohisto-chemical (IHC) staining intensity has been previously associated with a more aggressive disease [19,20], and we included this variable when we performed logistic regression predicting invasive disease, for each of the loci individually, dichotomizing the methylation extent at the median, and controlling for other potential confounders. In models controlled for age, gender, and TP53 immunohistochemical staining intensity of the tumor, only HOXB2 promoter methylation demonstrated an independent association with invasive tumors (Table 2). Tumors with HOXB2 methylation had an OR of 7.7 (95% CI 3.3, 18.2) for being an

Table 1. Demographics of the subject populations.

Characteristics	Series I – Profiled on Array (n = 73)	Series 2 – Profiled on Array (n = 264)	Series 3 – Used for Confirmation (n = 245)
Age, mean ±SD	62.1±9.3	65.1±9.9	62.9±9.1
Gender, n (%)			
Male	53 (73%)	200 (76%)	190 (78%)
Female	20 (27%)	64 (24%)	55 (22%)
Invasive Stage Tumor, n (%)*			
Non-invasive	42 (58%)	189 (72%)	181 (74%)
Low Grade (1–2)	35 (83%)	162 (86%)	158 (87%)
High Grade (3)	7 (17%)	27 (14%)	23 (13%)
Invasive	31 (42%)	75 (28%)	64 (26%)
Tumor TP53 IHC Staining Intensity, n (%)*			
Low (1,2)	50 (68%)	195 (74%)	189 (77%)
High (3+)	23 (32%)	69 (26%)	55 (23%)

*1 tumor in series 1 and one in the confirmation series did not have TP53 IHC staining data.
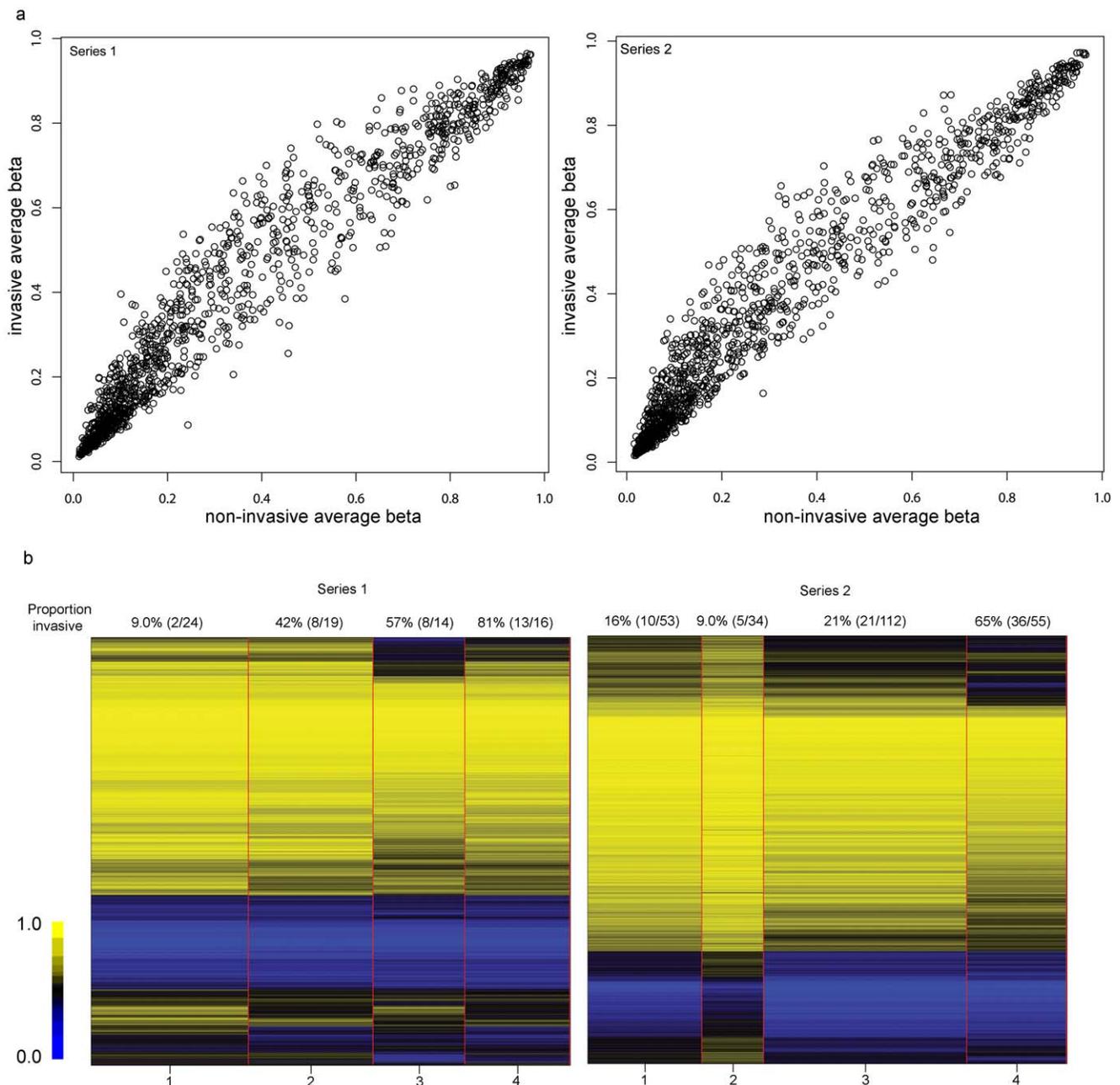doi:10.1371/journal.pone.0012334.t001

Figure 1. Methylation bead array profiles distinguish non-invasive and invasive bladder tumors. (A) Scatterplot of average methylation beta values in non-invasive bladder tumors (x-axis) and average methylation beta values in invasive tumors (y-axis) in series 1 (n = 73) and series 2 (n = 264) samples. (B) Recursive partition mixture models of each series result in 4 classes separated by red vertical lines, with the width of the classes corresponding to the number of samples in each class. Loci are represented as rows with the mean methylation for the class depicted. Above each class is the prevalence of invasive tumors (n invasive/total n) within the class. Methylation profiles significantly distinguish bladder tumor from non-diseased bladder epithelium (P<0.00001).
doi:10.1371/journal.pone.0012334.g001

invasive tumor. This result held up in a model controlling for all 3 loci as well as TP53 IHC staining intensity and patient age and gender, where *HOXB2* methylation was associated with an 8.6 fold increased risk of being an invasive tumor (95% CI 3.4, 21.7).

Methylation of HOXB2, FRZB, and KRT13 are Associated with Aggressive Non-invasive Bladder Cancer

As we are interested in determining how methylation contributes to progression and aggressiveness of tumors, we examined the association between methylation extent and tumor grade within the non-invasive tumors. Figure 4 demonstrates that there is significantly greater extent of methylation of *HOXB2* (P<0.01), *KRT13* (P<0.01) and *KRT13* (P = 0.0001) in high grade (3) compared to low grade (1,2) non-invasive tumors. To control for potential confounding, we performed multivariable logistic regression analysis to examine the association between gene promoter methylation extent and tumor grade. Initially, we modeled each gene promoter individually, and observed signifi-
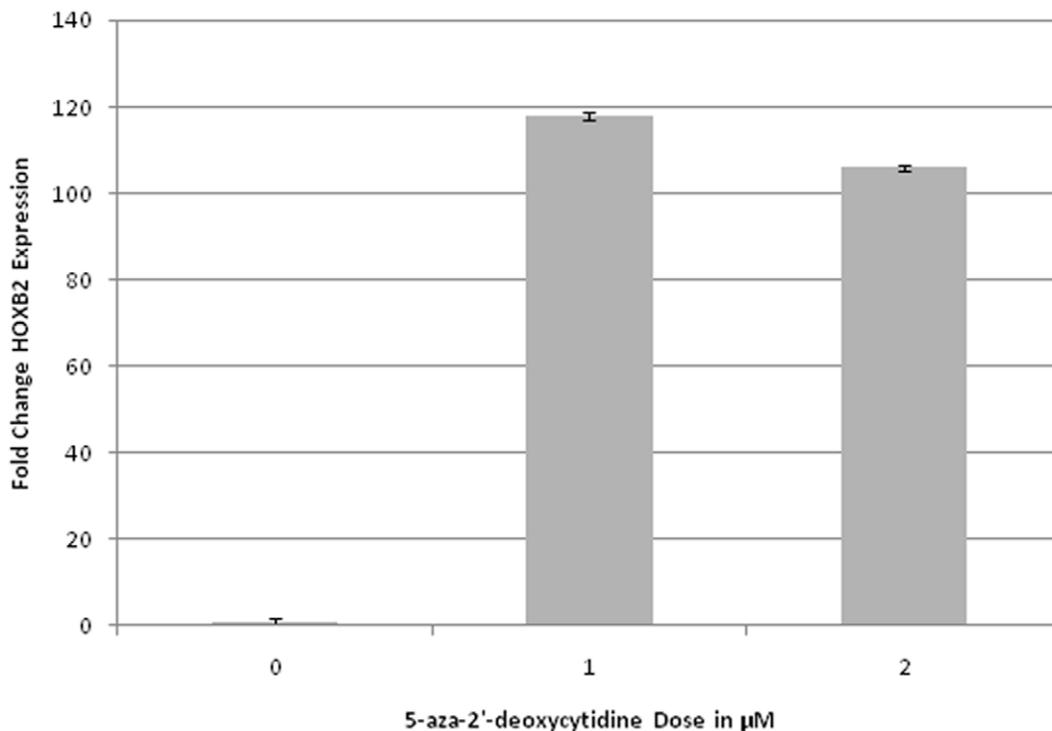
Figure 2. HOXB2 is re-expressed following treatment of HTB-9 bladder cancer cells with 5-aza-2′deoxycytidine. Quantitative RT-PCR analysis was used to determine the gene expression of *HOXB2* in human bladder carcinoma cell line HTB-9, following a 5-day treatment with 1 or 2 µM 5-aza-2′-deoxycytidine or mock treatment. Bars represent the mean fold change in expression compared to mock treatment for 4 or 6 replicate experiments, and error bars denote the standard errors.
doi:10.1371/journal.pone.0012334.g002

cant associations between *FRZB* (OR 2.9, 95% CI 1.1, 7.9), and *KRT13* (OR 3.3, 95% CI 1.1, 10.1), but only a borderline significant association between *HOXB2* methylation and high tumor grade (OR 2.6, 95% CI 0.9, 6.9, Table S2). We then examined the additive effects of methylation of these 3 loci by modeling the association between methylation of all three loci compared to having none, 1, or 2 methylated and high-grade non-invasive disease (Table 3). Controlled for age, gender, and TP53 IHC staining intensity, methylation of all 3 loci was associated with a significant 7.4 fold increased risk of being a high grade tumor (95% CI 2.5, 22.1).

Discussion

Compared to gene-expression based markers, DNA methylation patterns are more stable and can be detected using various approaches requiring relatively small amounts of patient materials. In addition, these markers can be objectively identified and quantified, providing potential improvement to reliability in histological subjective diagnoses such as tumor grades or compared to immunohistochemical staining patterns which rely on individual interpretation of intensity and location of staining. For example, the markers identified and independently validated in this study hold promise as biomarkers useful in bladder cancer screening and follow-up for recurrence and progression and could be applied and tested for clinical utility in easy to collect urine sediments [9,10,11,13,21]. In addition, discovery of genes and pathways epigenetically altered in bladder tumors suggest novel targets for therapeutic intervention, especially when considering the poor prognosis associated with invasive bladder cancers. Identification of important and useful DNA methylation based

markers is not without challenge, particularly as there is often a high degree of correlation between these alterations [22]. Therefore appropriate statistical approaches must be utilized to examine these alterations and to identify those alterations with the greatest clinical utility. We believe that our 2 stage approach provides for the most robust and generalizable identification of DNA methylation alterations which can be further examined in prospective studies. Using several different statistical approaches, our selection strategy for informative markers allowed us to identify a manageable number of potentially useful and robust biomarkers for laboratory validation as well as replication in independent samples. Most importantly, our validation of these markers in an independent group of patients provides compelling evidence of their potential utility and biological importance.

Previous studies examining panels of candidate genes have identified specific genes and panels of loci which are predictive of both the presence of tumor when examined in urine sediment [23] and the progression from non-invasive to invasive disease [15,24,25,26] when examined within the tumor tissue. Specific genes such as *RASSF1A*, *RARB*, *BAMBI* and the *SFRP* family have been associated with higher grade and higher stage bladder cancer [8,27,28]. Our strategy did not identify these specific genes, and this discrepancy may be due to a number of differences between our approach and those previously utilized. We have made use of a population-based case series sample, while many previous studies have utilized hospital-based convenience series, and thus may have had some bias introduced in the samples examined. The array platform used in this study is limited in that it examines only a subset of all genes and only 1 or 2 specific CpG sites for those genes, and thus may miss the specific genes of loci previously described. Finally, our analytical approach, as it is based on the
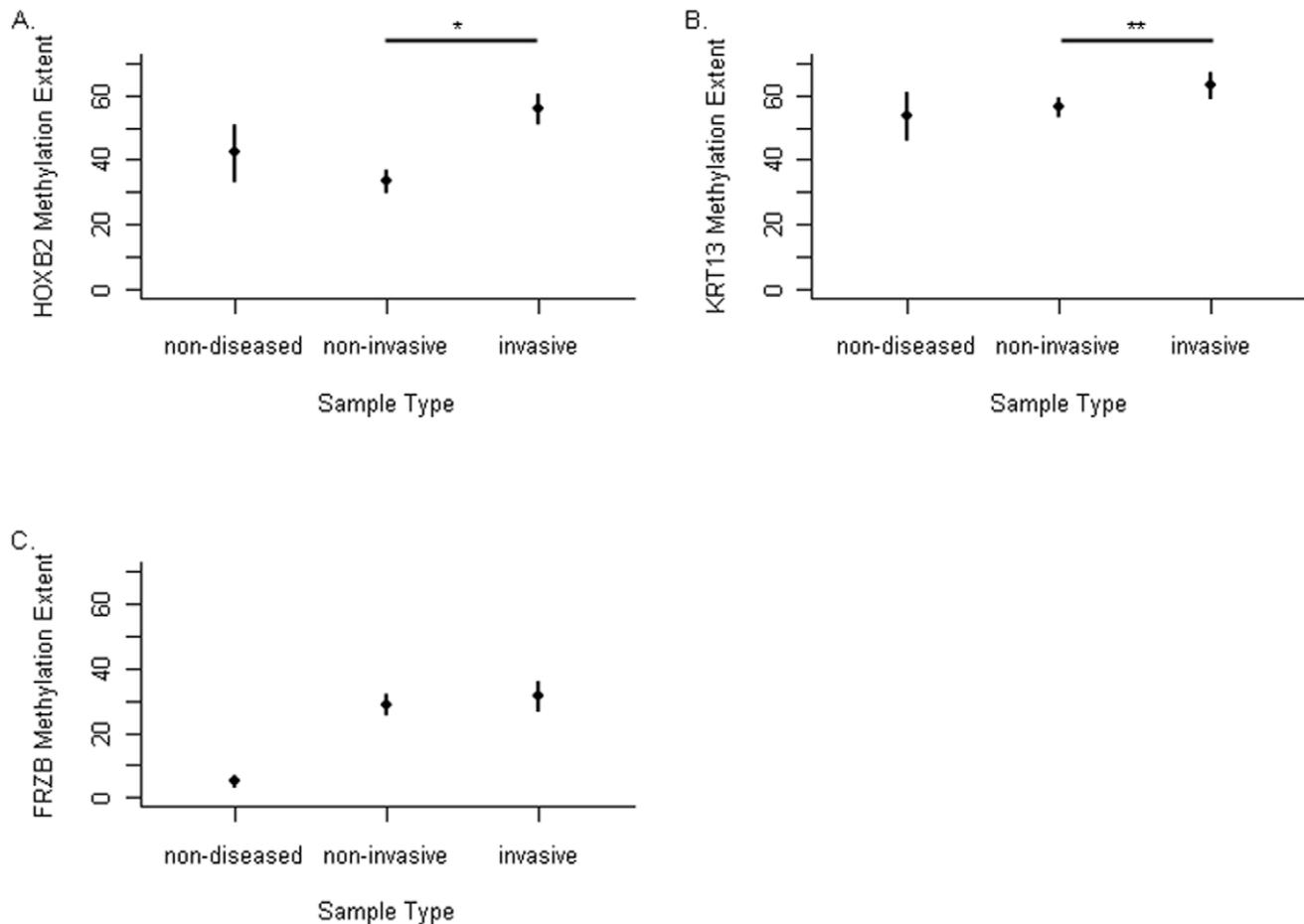
Figure 3. Bisulfite pyrosequencing reveals a greater extent of methylation in invasive bladder cancer. The mean (filled circle) and 95% confidence intervals of the extent of methylation of the promoter regions of (A) *FRZB*, (B) *KRT13*, and (C) *HOXB2* are depicted (y-axis) comparing non-diseased bladder epithelium, non-invasive tumors and invasive tumors (x-axis). The differences in the extent of methylation across these groups is significantly different (Kruskal-Wallis test), for *HOXB2* (P<0.00001) and *KRT13* (P<0.04), and *FRZB* (P<0.003). Comparisons specifically of noninvasive compared to invasive disease also revealed significant differences for *HOXB2* *P<0.00001 and *KRT13* **P<0.02.
doi:10.1371/journal.pone.0012334.g003

identification of panels of genes may identify different loci than those used in candidate analyses, but hypermethylation of those previously described candidates, may, in fact, be highly correlated to the methylation patterns identified herein. Further study combining previously identified loci with our own new targets in a prospective fashion is needed to more definitively determine those loci with the greatest clinical utility.

The biologic underpinnings of non-invasive yet recurrent disease versus invasive disease remain to be fully elucidated [29]. Our findings indicate that there are large differences in methylation between invasive and non-invasive tumors, with invasive tumors exhibiting a general increase in methylation compared to non-invasive tumors. As DNA methylation at any one CpG is a binary measure (either the cytosine is methylated or not) the difference in the degree of methylation may reflect potentially greater homogeneity in the selection of epigenetically altered cells in the development of invasive disease, a result consistent with our previous work examining only a limited panel of gene promoters in a single series of bladder tumors [14]. This greater degree of methylation may be driving the features of the invasive phenotype, or alternatively may be a consequence of selective pressures related to this phenotype. Although our approach in examining incident disease cannot fully distinguish

these possibilities, it provides data illustrative of the urgent need for further examination of this epigenetic phenomenon.

Identification of specific genes whose DNA methylation pattern is associated with tumor phenotypes, including stage and grade also illuminates the particularly important biological processes and pathways required for genesis of these histologically defined tissue states. The *FRZB* gene (also known as *SFRP3*) is a member of the secreted frizzled receptor family of soluble proteins which binds to and antagonizes the WNT receptor. We and others have previously shown that hypermethylation of *SFRP* genes is strongly and significantly associated with invasive bladder cancer, confirming the importance of this pathway in this phenotype [8,12]. Of the *SFRP* genes, only *SFRP1*, (in addition to *FRZB*) is profiled on the array, and at a different CpG location than the primers used in our and previous reports of its methylation [8,30]. This may explain why this loci was not identified in our screening approach. Additional, more genomically dense investigation of the WNT pathway is warranted, as this may suggest a novel route for therapeutic intervention for this disease.

HOXB2 is a member of the homeobox family of transcription factors, and is encoded on chromosome 17 as part of a gene cluster with other HOX family members. Interestingly, we observe that non-diseased bladder epithelium exhibits a similar extent of

Table 2. Individual gene methylation associations with invasive bladder cancer.

	Non-invasive n (%)	Invasive n (%)	Invasive OR (95% CI)*	Non-invasive n (%)	Invasive n (%)	Invasive Disease OR (95% CI)*	Non-invasive n (%)	Invasive n (%)	Invasive Disease OR (95% CI)*	Non-invasive n (%)	Invasive n (%)	Invasive Disease OR (95% CI)†
Total N	171	61		170	62		177	61		162	57	
HOXB2												
Negative (≤median)	110 (93)	8 (7)	1.0 (referent)	–	–	–	–	–	–	104 (94)	7 (6)	1.0 (referent)
Positive (>median)	61 (54)	53 (46)	7.7 (3.3, 18.2)	–	–	–	–	–	–	58 (54)	50 (46)	8.6 (3.4, 21.7)
KRT13												
Negative (≤median)	–	–	–	91 (80)	22 (20)	1.0 (referent)	–	–	–	86 (82)	19 (18)	1.0 (referent)
Positive (>median)	–	–	–	79 (66)	40 (34)	1.3 (0.7, 2.7)	–	–	–	76 (67)	38 (33)	1.0 (0.4, 2.3)
FRZB												
Negative (≤median)	–	–	–	–	–	–	93 (80)	24 (20)	1.0 (referent)	84 (78)	23 (22)	1.0 (referent)
Positive (>median)	–	–	–	–	–	–	84 (69)	37 (31)	1.4 (0.7, 2.7)	78 (70)	34 (30)	0.9 (0.4, 2.0)
TP53 IHC Staining Intensity												
Low (<3)	154 (86)	26 (14)	1.0 (referent)	154 (86)	26 (14)	1.0 (referent)	160 (86)	26 (14)	1.0 (referent)	146 (85)	26 (15)	1.0 (referent)
High (3+)	17 (33)	35 (67)	7.4 (3.4, 16.0)	16 (31)	36 (69)	12.4 (5.9, 25.9)	17 (33)	35 (67)	12.1 (5.9, 24.9)	16 (34)	31 (66)	6.1 (2.7, 13.8)

*Model is adjusted for age, gender, and TP53 Staining Intensity.
†Model is adjusted for age, gender, each gene methylation variable, and TP53 Staining Intensity and includes only those subjects with data for all covariates.
doi:10.1371/journal.pone.0012334.t002

methylation at *HOXB2* and *KRT13* as the non-invasive disease, while at *FRZB* the non-diseased tissue has a significantly lower extent of methylation than either of the 2 tumor types. We also observe, in the human bladder carcinoma cell line HTB-9, which demonstrates an extent of methylation similar to that observed in invasive tumors, that treatment with the methylation inhibitor 5-aza-2′-deoxycytidine leads to increased expression of *HOXB2*, suggesting that methylation of this gene may be functionally leading to the inactivation of this gene. The results in *HOXB2* are consistent with the literature which has described a bivalent domain structure of the chromatin of homeobox genes, wherein both pluripotent and terminally differentiated tissue exhibit marks of both active and repressive chromatin within the same region [31,32], and thus an intermediate extent of DNA methylation. Invasive bladder cancers appear targeted for hypermethylation of this region, a phenomenon previously reported across polycomb group controlled genes in colorectal cancer cells [33]. Such a differential alteration of a key developmental gene may be critical in defining the phenotypes of these tumors and explain the behavior and outcome from these disparate forms of this disease.

KRT13 encodes cytokeratin-13 which has been shown to have specific expression in cervical squamous tumors and in mucinous cervical type adenocarcinomas [34]; to our knowledge, epigenetic alterations to this gene have not been reported. Epigenetic alteration of this gene may affect its expression pattern, and such changes may be signs of a more de-differentiated phenotype, considered characteristic of aggressive, high grade tumors.

We demonstrated in an independent confirmation series of tumors that the extent of methylation of *HOXB2* and *FRZB* is associated with tumor grade, independent of each other and of TP53 protein immunohistochemical staining. TP53 staining has previously been examined and has been touted as a useful prognostic marker in this disease [35,36,37], although there is

conflicting data on its utility [36,38,39,40]. A recent meta-analysis examining its utility suggests that there is not appropriate evidence to suggest utilizing this marker clinically [20], as it may provide no addition prognostic information other than a strong association with the invasive phenotype of the disease [41,42]. We cannot yet suggest that the markers discovered using our approach hold any more clinical utility than pathological grading and staging of tumors, but we believe that our data suggest specific cellular pathways that are disrupted in this disease. Thus, these pathways could be explored intensively as novel targets for therapeutic strategies. In addition, we suggest that future work, examining the utility of these markers in a prospective fashion is needed, as they may be useful in determining which tumors may progress. Indeed, these markers might also contribute to clinical efforts to follow patients with less invasive methods, to determine if their tumor has recurred or progressed, thereby requiring more aggressive treatment approaches.

Methods

Study Population and Sample Ascertainment

All study participants provided written informed consent under the approval of the institutional review boards of Dartmouth Medical School and Brown University. We utilized two, independent, non-consecutive population-based series of bladder cancer cases. The first, consisting of tumors from 344 individuals involved in a case-control study of incident bladder cancer in New Hampshire, diagnosed between July 1994 and June 1998 [43], and the second consisting of tumors from 264 individuals diagnosed between January 1, 2002 to July 30, 2004 [44]. Although separate in time and scope, these two studies utilized identical recruitment procedures and study personnel as well as identical protocols for the ascertainment of pathology materials for molecular examina-
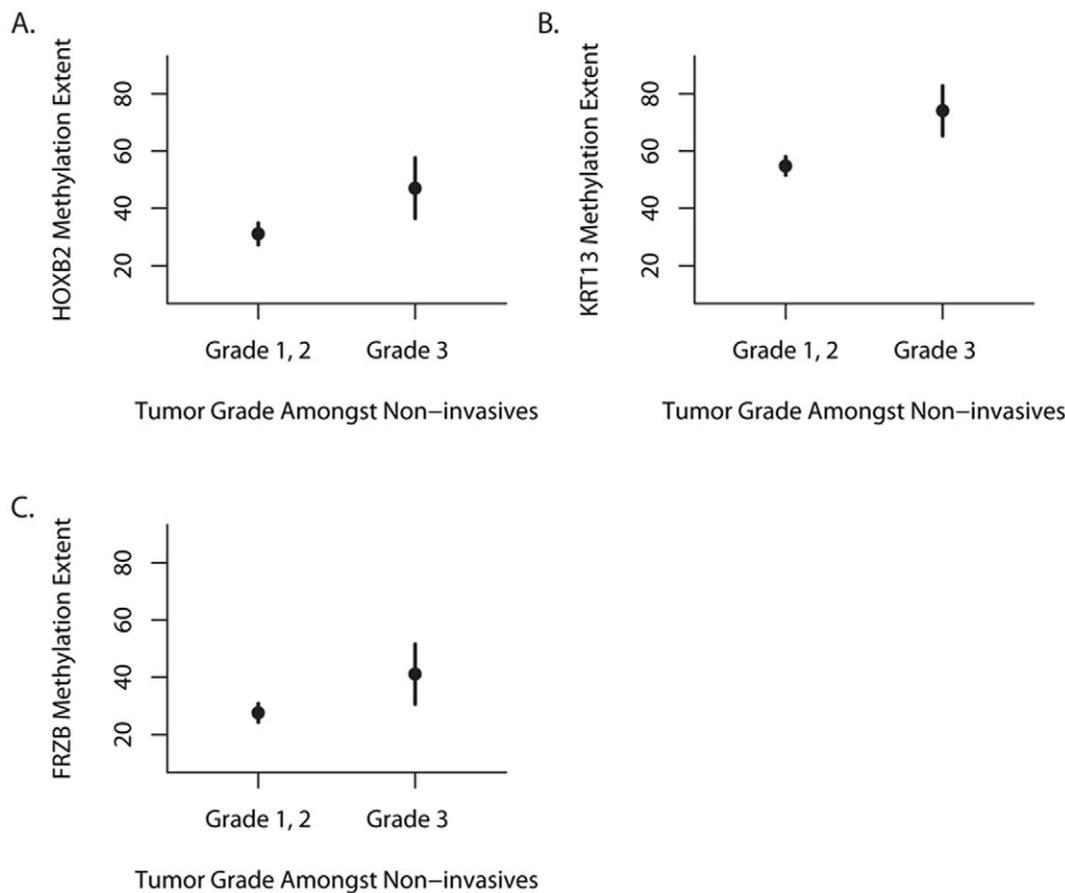
Figure 4. Bisulfite pyrosequencing demonstrates that extent of methylation is associated with tumor grade in non-invasive tumors. The mean (filled circle) and 95% confidence intervals of the extent of methylation of the promoter regions of (A) *FRZB*, (B) *KRT13*, and (C) *HOXB2* are depicted (y-axis) comparing low (1,2) to high (3+) grade tumors (x-axis). P-values resulting from the Wilcoxon Rank Sums demonstrate that these differences are significant for *HOXB2* (P<0.01), *KRT13* (P<0.01) and *FRZB* (P = 0.0001).
doi:10.1371/journal.pone.0012334.g004

tions. Bladder tumors from both series were reviewed by the study pathologist (A.R.S.) and classified according to the 1973 and 2004 World Health Organization guidelines for bladder tumors. The study pathologist identified the appropriate block from which the

Table 3. Methylation of *HOXB2*, *FRZB* and *KRT13* is independently associated with high grade disease in non-invasive bladder cancer.

	Grade 1, 2	Grade 3	Grade 3 OR
	n(%)	**n(%)**	**(95% CI)**
Total N	143	19	
HOXB2, FRZB, and KRT13 Methylation >Median			
None, 1, or 2 Genes	125 (94)	8 (6.0)	1.0 (referent)
All 3 Genes	18 (62)	11 (38)	7.4 (2.5, 22.1)
TP53 IHC Staining Intensity			
Low (<3)	133 (91)	13 (9.0)	1.0 (referent)
High (3+)	10 (62)	6 (38)	4.7 (1.3, 17.4)

Each model is controlled for variables in columns, as well as age and gender.
doi:10.1371/journal.pone.0012334.t003

tumor samples used in these analyses were obtained, and the proportion of malignant cells in each sample was estimated. All samples used in the examination contained >75% tumor sample. Carcinoma in-situ was excluded from analysis due to limited sample size. Table 1 describes the characteristics of the subjects included in the final analysis. In addition, four non-diseased bladder epithelium samples were obtained from the National Disease Research Interchange (NDRI) all from autopsy specimens of individuals who did not have a diagnosis of cancer.

DNA Extraction and Methylation Analysis

Tumor sections with the greatest proportion of malignant tissue were selected by the study pathologist for use in our molecular analyses. DNA was extracted and sodium bisulfite modified following standard procedures as described in Marsit, et al. [18]. A total of 82 tumors from the first series and all 264 tumors from the second series were profiled for the methylation status at 1505 CpG loci using the Illumina GoldenGate® methylation bead arrays. Bead arrays were run at the UCSF Institute for Human Genetics, Genomics Core Facility according to the manufacturer's protocol and as described in Bibikova, et al. [45].

Statistical analysis

We assembled data with BeadStudio Methylation software from the array manufacturer (Illumina, San Diego, CA) and quality

assessment performed as in Christensen et al. [16] resulting in seven samples (9%), where >75% of loci had a detection p-value $>1*10^{-5}$, being dropped from analysis. A similar quality control for CpG loci eliminated those loci with median detection p-value >0.05 (n = 8, 0.5%).

Subsequent analyses were carried out in The R Package. For exploratory/visualization purposes, hierarchical clustering using the Manhattan metric and average linkage was performed. For locus by locus analyses, associations at individual CpG loci were tested with a generalized linear model (GLM). Under the assumption that the average beta values follow a beta distribution [46], a quasi-binomial model (logit link, binomial variance, and non-unit scale parameter) was used; note that the logit link function imposes an appropriate constraint on the mean average beta. To adjust for multiple comparisons in scanning the 1413 autosomal CpG loci, P-values for associations between average beta and invasive disease used false discovery rate correction and q-values computed by the *qvalue* package in R. An FDR q-value of <0.05 was considered significant for this examination.

Random Forests [47](http://www.math.usu.edu/~adele/forests/), was used to build classifiers of invasive stage (versus noninvasive) by CpG average beta values. Analysis was conducted in R using the *randomForest* package (version 4.5-18) by Liaw and Wiener. At each node of the tree, a random sample of m out of the total M variables was chosen and the best split is found among the m variables. The default value for m in the Random Forest R package is $\sqrt{M}$. In this analysis we tested a range of m from half of $\sqrt{M}$ to two times $\sqrt{M}$ and used m that gave the lowest prediction error; in this case $m = 38$. The OOB error rate is the percentage of time the RF prediction is incorrect. A test for association between methylation (predictors) and sample type was conducted by comparing the OOB obtained on the data set with the null distribution of OOB errors obtained by permuting sample type labels and running the RF procedure 100 times. We also used the second series of 264 tumors as a validation set, computing the error rate as each tree is constructed, to obtain an estimated error rate independent of the first series. Finally, we used the variable importance scores, percent change in mean squared error (MSE), to identify loci that had the greatest influence on classification, by choosing those loci whose percent change in MSE was greater than 5%.

For inference, data were clustered using a mixture model [48] with a mixture of beta distributions [49], and the number of classes was determined by Bayesian information criterion (BIC) [50,51]. The mixture model was fit by recursively partitioning the data using a 2-class mixture model, with a variant of BIC used as a criterion for the split, as described in [46]. Class membership was obtained from the mixture model using an empirical Bayes procedure, and subsequent associations with invasive stage were tested via permutation test with 10,000 permutations each using the standard chi-square goodness-of-fit test. The output of mixture models includes a class-specific beta distribution F_{jk} for each locus j and class k. These distributions imply a receiver operating curve (ROC) for distinguishing two classes k and l by the relation $ROC_{jkl}(r) = F_{jk}(F_{jl}^{-1}(r))$ or $ROC_{jkl}(r) = 1 - F_{jk}(F_{jl}^{-1}(1-r))$, whichever curve lies above the identity line. In either case, the area-under-the-ROC (AUC), computed as $AUC_{jkl} = \left| \frac{1}{2} - \int_0^1 F_{jk}(F_{jl}^{-1}(r))dr \right| + \frac{1}{2}$, which can be used to determine the influence of locus j on the distinction between classes k and l. A similar procedure can be used to distinguish class k from the others: we use a normal approximation G_{jk} for the distribution of the combined classes other than k, and compute

$AUC_{jkl} = \left| \frac{1}{2} - \int_0^1 G_{jk}(F_{jk}^{-1}(r))dr \right| + \frac{1}{2}$. We then ranked each locus j by its computed AUC_{jk}, and selected those with AUC>0.90.

To identify loci for subsequent validation as predictors of invasive vs. non-invasive disease, we computed the intersection of the loci that were most significant (Q<0.05) in the locus-by-locus analysis, the top ranking loci from the RF analysis (% change in MSE>6%), and the loci having AUC>0.75 for distinguishing the class of interest compared to others in each tumor series, and then identified those loci overlapping in both series.

Bisulfite Pyrosequencing

Quantification of cytosine percent methylation was performed by pyrosequencing bisulfite-converted DNA using the PyroMark MD pyrosequencing system (Qiagen, Valencia, CA). Specific pyrosequencing primers were designed to amplify array CpG sites and as many downstream CpGs as conditions permitted (1 to 5 additional) using Biotage Assay Design Software v1.0.6 (Qiagen). Assays designs were attempted for all 5 of the loci identified as overlapping from the 3 analysis strategies and assays were successfully designed for 4 of the loci. Primer details and the genomic location of the region sequenced relative to the transcription start site are provided in Table S3. The sequences used were based on the reference sequence assembly 36.1. All PCR reactions, performed using Qiagen Hot Star Taq polymerase, included a no template control, unmodified DNA control, and 7 standardized percent methylation controls (0%, 15%, 25%, 45%, 65%, 75%, and 100%) derived from Qiagen EpiTect® PCR control DNA set samples. Sequencing reactions used 10µl of PCR product and were run according to instrument/manufacturer protocols (Biotage).

Statistical Analysis of Pyrosequencing Data

Associations between pyrosequencing percent methylation values at individual CpG sites were examined using the Spearman rank correlation. The mean value of CpG methylation across the examined CpG sites was used in further analyses. To examine the association between methylation extent as a continuous variable and tumor grade, we utilized the non-parametric Kruskal-Wallis one-way analysis of variance. To adjust for potential confounders and to determine the odds ratio (OR) and 95% confidence interval (CI) for high grade disease, we used multivariable unconditional logistic regression analysis to examine the association between methylation extent (in quartiles) at individual sites and tumor grade (grade 1 vs. grade 2, 3, 4), with adjustment for potential confounders. We also included a term to control for percent of tumor cells within sample in the model, and found that there were no significant changes in the effect estimates or significance of the main effects, and thus this term was not considered a confounder as was removed from the model.

Cell Culture and Drug Treatment

Human bladder carcinoma cell lines UMUC3 and HTB-9 were purchased from the American Type Culture Collection (Manassas, VA). Cell lines were grown in their respective vendor-recommended culture media at 37°C and 5% CO_2 and passaged every 3–5 days. Approximately 1 million cells were harvested from each and DNA isolated and sodium bisulfite modified as above to examine methylation status of the candidate loci. 5-aza-2'-deoxycytidine was obtained from Sigma-Aldrich (St. Louis, MO), and was prepared in sterile DMSO. The concentrations used for the experiment were chosen to mimic those observed clinically in human plasma [52] and that did not result in significant loss in cell viability. For experiments, cells were seeded 24 hours before drug treatment, 20,000 cells per well into a 6-well cell culture plate. The media was replaced with that containing 5-

aza-2′-deoxycytidine, and was changed on day 1 and 3. Cells were harvested on day 5 and immediately subjected to RNA isolation.

RNA Isolation and Quantitative RT-PCR Analysis

Total RNA was isolated from cells using the Qiagen RNEasy RNA isolation system, following manufacturer's protocols, and including an on-column DNAse digestion. The cDNA was prepared from 150 ng total RNA using the Taqman Gene Expression system (Applied Biosystems, Carlsbad, CA). cDNA, diluted 1:5 was subjected to RT-PCR for HOXB2 (Hs01911157_s1) and ACTB (hu ACTB) as the referent gene using pre-designed Taqman primers and probes on an ABI 7500 Fast Real Time PCR system (Applied Biosystems). All RT-PCRs were run in triplicate and no-template and no reverse transcription controls were included for each assay. Fold change was calculated using the Pfaffl method [53].

Supporting Information

Figure S1 Diagram of the methodology used in the selection of loci for follow-up analyses and validation.
Found at: doi:10.1371/journal.pone.0012334.s001 (0.27 MB TIF)

Figure S2 Confirmation of DNA methylation of FRZB (A), STAT5A (B), KRT13 (C), and HOXB2 (D) by pyrosequencing in samples analyzed on the array. Dots represent percent methylation at each position examined (POS) in non-invasive (N) compared to invasive (I) tumor tissue. The gray box encloses the position which represents that position examined on the array. *P<0.05, **P<0.0001.
Found at: doi:10.1371/journal.pone.0012334.s002 (1.68 MB TIF)

Table S1 Results of locus by locus analysis comparing invasive to non-invasive bladder cancers with Q<0.001.
Found at: doi:10.1371/journal.pone.0012334.s003 (0.56 MB PDF)

Table S2 Individual gene promoter methylation is associated with high grade in non-invasive bladder cancer.
Found at: doi:10.1371/journal.pone.0012334.s004 (0.05 MB PDF)

Table S3 Primer sequences used for bisulfite pyrosequencing reactions.
Found at: doi:10.1371/journal.pone.0012334.s005 (0.08 MB PDF)

Author Contributions

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, et al. (2009) Cancer statistics, 2009. CA Cancer J Clin 59: 225–249.
2. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R (2003) The health economics of bladder cancer - A comprehensive review of the published literature. Pharmacoeconomics 21: 1315–1330.
3. Ockrim JL, Abel PD (2007) Treatment Options in Superficial (pTa/pT1/CIS)/ Bladder Cancer. In: Waxman J, ed. Urological cancers in clinical practice. London: Springer. pp 75–101.
4. Iacobuzio-Donahue CA (2009) Epigenetic changes in cancer. Annu Rev Pathol 4: 229–249.
5. Catto JW, Miah S, Owen HC, Bryant H, Myers K, et al. (2009) Distinct microRNA alterations characterize high- and low-grade bladder cancer. Cancer Res 69: 8472–8481.
6. Mathews LA, Crea F, Farrar WL (2009) Epigenetic gene regulation in stem cells and correlation to cancer. Differentiation 78: 1–17.
7. Esteller M (2003) Cancer epigenetics: DNA methylation and chromatin alterations in human cancer. Adv Exp Med Biol 532: 39–49.
8. Marsit CJ, Karagas MR, Andrew A, Liu M, Danaee H, et al. (2005) Epigenetic inactivation of SFRP genes and TP53 alteration act jointly as markers of invasive bladder cancer. Cancer Res 65: 7081–7085.
9. Yu J, Zhu T, Wang Z, Zhang H, Qian Z, et al. (2007) A novel set of DNA methylation markers in urine sediments for sensitive/specific detection of bladder cancer. Clin Cancer Res 13: 7296–7304.
10. Hoque MO, Begum S, Topaloglu O, Chatterjee A, Rosenbaum E, et al. (2006) Quantitation of promoter methylation of multiple genes in urine DNA and bladder cancer detection. J Natl Cancer Inst 98: 996–1004.
11. Sathyanarayana UG, Maruyama R, Padar A, Suzuki M, Bondaruk J, et al. (2004) Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes. Cancer Res 64: 1425–1430.
12. Stoehr R, Wissmann C, Suzuki H, Knuechel R, Krieg RC, et al. (2004) Deletions of chromosome 8p and loss of sFRP1 expression are progression markers of papillary bladder cancer. Lab Invest 84: 465–478.
13. Urakami S, Shiina H, Enokida H, Kawakami T, Kawamoto K, et al. (2006) Combination analysis of hypermethylated Wnt-antagonist family genes as a novel epigenetic biomarker panel for bladder cancer detection. Clin Cancer Res 12: 2109–2116.
14. Marsit CJ, Houseman EA, Schned AR, Karagas MR, Kelsey KT (2007) Promoter hypermethylation is associated with current smoking, age, gender and survival in bladder cancer. Carcinogenesis 28: 1745–1751.
15. Aleman A, Adrien L, Lopez-Serra L, Cordon-Cardo C, Esteller M, et al. (2008) Identification of DNA hypermethylation of SOX9 in association with bladder cancer progression using CpG microarrays. Br J Cancer 98: 466–473.
16. Christensen BC, Houseman EA, Godleski JJ, Marsit CJ, Longacker JL, et al. (2009) Epigenetic profiles distinguish pleural mesothelioma from normal pleura and predict lung asbestos burden and clinical outcome. Cancer Res 69: 227–234.
17. Killian JK, Bilke S, Davis S, Walker RL, Killian MS, et al. (2009) Large-scale profiling of archival lymph nodes reveals pervasive remodeling of the follicular lymphoma methylome. Cancer Res 69: 758–764.
18. Marsit CJ, Christensen BC, Houseman EA, Karagas MR, Wrensch MR, et al. (2009) Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. Carcinogenesis 30: 416–422.
19. Kelsey KT, Hirao T, Hirao S, Devi-Ashok T, Nelson HH, et al. (2005) TP53 alterations and patterns of carcinogen exposure in a U.S. population-based study of bladder cancer. Int J Cancer 117: 370–375.
20. Malats N, Bustos A, Nascimento CM, Fernandez F, Rivas M, et al. (2005) P53 as a prognostic marker for bladder cancer: a meta-analysis and review. Lancet Oncol 6: 678–686.
21. Kim YK, Kim WJ (2009) Epigenetic markers as promising prognosticators for bladder cancer. Int J Urol 16: 17–22.
22. Marsit CJ, Houseman EA, Christensen BC, Eddy K, Bueno R, et al. (2006) Examination of a CpG island methylator phenotype and implications of methylation profiles in solid tumors. Cancer Res 66: 10621–10629.
23. Yates DR, Rehman I, Meuth M, Cross SS, Hamdy FC, et al. (2006) Methylational urinalysis: a prospective study of bladder cancer patients and age stratified benign controls. Oncogene 25: 1984–1988.
24. Yates DR, Rehman I, Abbod MF, Meuth M, Cross SS, et al. (2007) Promoter hypermethylation identifies progression risk in bladder cancer. Clin Cancer Res 13: 2046–2053.
25. Aleman A, Cebrian V, Alvarez M, Lopez V, Orenes E, et al. (2008) Identification of PMF1 methylation in association with bladder cancer progression. Clin Cancer Res 14: 8236–8243.
26. Kim EJ, Kim YJ, Jeong P, Ha YS, Bae SC, et al. (2008) Methylation of the RUNX3 promoter as a potential prognostic marker for bladder tumor. J Urol 180: 1141–1145.
27. Jarmalaite S, Jankevicius F, Kurgonaite K, Suziedelis K, Mutanen P, et al. (2008) Promoter hypermethylation in tumour suppressor genes shows association with stage, grade and invasiveness of bladder cancer. Oncology 75: 145–151.
28. Khin SS, Kitazawa R, Win N, Aye TT, Mori K, et al. (2009) BAMBI gene is epigenetically silenced in subset of high-grade bladder cancer. Int J Cancer 125: 328–338.
29. Dinney CP, McConkey DJ, Millikan RE, Wu X, Bar-Eli M, et al. (2004) Focus on bladder cancer. Cancer Cell 6: 111–116.
30. Suzuki H, Watkins DN, Jair KW, Schuebel KE, Markowitz SD, et al. (2004) Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. Nat Genet 36: 417–422.
31. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, et al. (2006) Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8: 532–538.
32. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315–326.

33. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, et al. (2007) Epigenetic stem cell signature in cancer. Nat Genet 39: 157–158.

34. Carrilho C, Alberto M, Buane L, David L (2004) Keratins 8, 10, 13, and 17 are useful markers in the diagnosis of human cervix carcinomas. Hum Pathol 35: 546–551.

35. Serth J, Kuczyk MA, Bokemeyer C, Hervatin C, Nafe R, et al. (1995) p53 immunohistochemistry as an independent prognostic factor for superficial transitional cell carcinoma of the bladder. Br J Cancer 71: 201–205.

36. Gardiner RA, Walsh MD, Allen V, Rahman S, Samaratunga ML, et al. (1994) Immunohistological expression of p53 in primary pT1 transitional cell bladder cancer in relation to tumour progression. Br J Urol 73: 526–532.

37. Watanabe R, Tomita Y, Nishiyama T, Tanikawa T, Sato S (1994) Correlation of p53 protein expression in human urothelial transitional cell cancers with malignant potential and patient survival. Int J Urol 1: 43–48.

38. Esuvaranathan K, Chiong E, Thamboo TP, Chan YH, Kamaraj R, et al. (2007) Predictive value of p53 and pRb expression in superficial bladder cancer patients treated with BCG and interferon-alpha. Cancer 109: 1097–1105.

39. Fleshner N, Kapusta L, Ezer D, Herschorn S, Klotz L (2000) p53 nuclear accumulation is not associated with decreased disease-free survival in patients with node positive transitional cell carcinoma of the bladder. J Urol 164: 1177–1182.

40. Grossman HB, Liebert M, Antelo M, Dinney CP, Hu SX, et al. (1998) p53 and RB expression predict progression in T1 bladder cancer. Clin Cancer Res 4: 829–834.

41. Hemal AK, Khaitan A, Dinda AK, Gupta NP, Seth A, et al. (2003) Prognostic significance of p53 nuclear overexpression in patients of muscle invasive urinary bladder carcinoma treated with cystectomy. Urol Int 70: 42–46.

42. Kelsey KT, Hirao T, Schned A, Hirao S, Devi-Ashok T, et al. (2004) A population-based study of immunohistochemical detection of p53 alteration in bladder cancer. Br J Cancer 90: 1572–1576.

43. Karagas MR, Tosteson TD, Blum J, Morris JS, Baron JA, et al. (1998) Design of an epidemiologic study of drinking water arsenic exposure and skin and bladder cancer risk in a U.S. population. Environ Health Perspect 106 Suppl 4: 1047–1050.

44. Baris D, Karagas MR, Verrill C, Johnson A, Andrew AS, et al. (2009) A case-control study of smoking and bladder cancer risk: emergent patterns over time. J Natl Cancer Inst 101: 1553–1561.

45. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, et al. (2006) High-throughput DNA methylation profiling using universal bead arrays. Genome Res 16: 383–393.

46. Houseman EA, Christensen BC, Yeh RF, Marsit CJ, Karagas MR, et al. (2008) Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions. BMC Bioinformatics 9: 365.

47. Breiman L (2001) Random Forests. Machine Learning 45: 5–32.

48. Siegmund KD, Laird PW, Laird-Offringa IA (2004) A comparison of cluster analysis methods using DNA methylation data. Bioinformatics 20: 1896–1904.

49. Ji Y, Wu C, Liu P, Wang J, Coombes KR (2005) Applications of beta-mixture models in bioinformatics. Bioinformatics 21: 2118–2122.

50. Fraley F, Raftery A (2002) Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc 97: 611–631.

51. Houseman EA, Coull BA, Betensky RA (2006) Feature-specific penalized latent class analysis for genomic data. Biometrics 62: 1062–1070.

52. Cashen AF, Shah AK, Todt L, Fisher N, DiPersio J (2008) Pharmacokinetics of decitabine administered as a 3-h infusion to patients with acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). Cancer Chemother Pharmacol 61: 759–766.

53. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29: e45.