Published in final edited form as:

Hum Genet. 2009 June ; 125(5-6): 527–539. doi:10.1007/s00439-009-0645-6.

Bladder cancer SNP panel predicts susceptibility and survival

Angeline S. Andrew and Jiang Gui

Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Dartmouth Medical School, 7927 Rubin 860, One Medical Center Drive, Lebanon, NH 03756, USA

Arthur C. Sanderson

Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Rebecca A. Mason and Elaine V. Morlock

Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Dartmouth Medical School, 7927 Rubin 860, One Medical Center Drive, Lebanon, NH 03756, USA

Alan R. Schned

Department of Pathology, Dartmouth Medical School, Lebanon, NH 03756, USA

Karl T. Kelsey and Carmen J. Marsit

Departments of Community Health, and of Pathology and Laboratory Medicine, Brown University, Providence, RI 02912, USA

Jason H. Moore and Margaret R. Karagas

Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Dartmouth Medical School, 7927 Rubin 860, One Medical Center Drive, Lebanon, NH 03756, USA

Abstract

Bladder cancer is the fourth most common malignancy in men and the eighth most common in women in western countries. Single nucleotide polymorphisms (SNPs) in genes that regulate telomere maintenance, mitosis, inflammation, and apoptosis have not been assessed extensively for this disease. Using a population-based study with 832 bladder cancer cases and 1,191 controls, we assessed genetic variation in relation to cancer susceptibility or survival. Findings included an increased risk associated with variants in the methyl-metabolism gene, *MTHFD2* (OR 1.7 95% CI 1.3–2.3), the telomerase *TEP1* (OR 1.8 95% CI 1.2–2.6) and decreased risk associated with the inflammatory response gene variant *IL8RB* (OR 0.6 95% CI 0.5–0.9) compared to wild-type. Shorter survival was associated with apoptotic gene variants, including *CASP9* (HR 1.8 95% CI 1.1–3.0). Variants in the detoxification gene *EPHX1* experienced longer survival (HR 0.4 (95% CI 0.2–0.8). These genes can now be assessed in multiple study populations to identify and validate SNPs appropriate for clinical use.

Introduction

In western countries, bladder cancer is the fourth most common malignancy in men and the eighth most common in women (Kirkali et al. 2005). The majority of this disease is attributed to cigarette smoking; bladder cancer risk is up to fourfold higher among cigarette smokers compared with non-smokers (Kirkali et al. 2005). A hereditary component is likely since a family history of bladder cancer and variations in genes that detoxify aromatic amines are

associated with increased risk (Garcia-Closas et al. 2005; Kantor et al. 1985; Murta-Nascimento et al. 2007b). We and others have also identified associated variations in DNA repair pathway genes, yet single nucleotide polymorphisms (SNPs) in important cancer regulatory pathways such as cell growth and apoptosis have not been comprehensively assessed for this disease (Andrew et al. 2006; Wu et al. 2006).

Bladder cancer generally carries a favorable prognosis, nevertheless, approximately 13,750 deaths occurred in the USA from bladder cancer in 2007 (Jemal et al. 2007). According to the US Statistics Epidemiology and End Results (SEER) program (1996–2002), the 5-year survival rate for localized disease is 94%, whereas for regional staged cancers the survival rate is only 47%. Those with distant metastasis at the time of diagnosis carry the lowest survival rate of 6% (Jemal et al. 2007). Established prognostic factors include tumor characteristics such as degree of invasion or carcinoma in situ (CIS) (Carroll et al. 2001). Higher grade of a tumor is also associated with a worse prognosis (Schned et al. 2007). Although these histopathologic factors can be used to predict prognosis, the behavior of each group is quite heterogeneous [reviewed in (Murta-Nascimento et al. 2007a)]. This inter-individual variation in survival highlights the potential for new genetic markers that can assist physicians in selecting appropriately aggressive treatment regimens.

We have assessed the role of genetic variation in bladder cancer risk and case survival using panel of SNPs in hypothesized cancer regulatory pathways, including cell cycle, cell growth, detoxification, telomerase, and apoptosis. Risk- or survival-associated single gene and multi-gene-combinations were ranked and tested using traditional multivariate statistical methods in a large, population-based case–control study in New Hampshire, USA.

Materials and methods

Study group

We identified all cases of bladder cancer among New Hampshire residents, ages 25–74 years from the State Cancer Registry. The study was conducted in two sequential phases based on cancer diagnosis date (Phase I 1 July 1994–30 June 1998, Phase II 1 July 1998–30 June 2001). Detailed methods have been described previously (Karagas et al. 1998). Briefly, we interviewed a total of 832 bladder cancer cases, which was 85% of the cases confirmed to be eligible for the study. Controls less than 65 years of age were selected using population lists obtained from the New Hampshire Department of Transportation. Controls 65 years of age and older were chosen from data files provided by the Centers for Medicare & Medicaid Services (CMS) of New Hampshire. For efficiency, we shared a control group with a study of non-melanoma skin cancer in New Hampshire covering an overlapping diagnostic period of 1 July 1993–30 June 1995 (Karagas et al. 1998). We selected additional controls for bladder cancer cases diagnosed from 1 July 1995 to 30 June 1998 frequency matched to these cases on age (25–34, 35–44, 45–54, 55–64, 65–69, 70–74 years) and gender. Most (>95%) of the subjects in this study are of Caucasian origin; and thus our analyses were not appreciably altered by restricting to Caucasians. We interviewed a total 1,191 controls (the total shared control group and additional controls), which was 70% of the controls confirmed to be eligible for the study.

Personal interview

Informed consent was obtained from each participant and all procedures and study materials were approved by the Committee for the Protection of Human Subjects at Dartmouth College. Consenting participants underwent a detailed in-person interview, usually at their home. Questions covered sociodemographic information (including level of education), lifestyle factors such as use of tobacco (including frequency, duration and intensity of smoking), family history of cancer, and medical history prior to the diagnosis date of the bladder cancer (cases)

or reference date assigned to controls. Recruitment procedures for both the shared controls from the non-melanoma skin cancer study and additional controls were identical and ongoing concomitantly with the case interviews. Case-control status and the main objectives of the study were not disclosed to the interviewers. To ensure consistent quality of the study interviewer, interviews were tape recorded with the consent of the participants and routinely monitored by the interviewer supervisor. To assess comparability of cases and controls, we asked subjects if they currently held a driver's license or a Medicare enrollment card. Subjects were asked to provide a blood sample (buccal sample was requested in the case of a refusal). Samples were maintained at 4°C and sent via courier to the study laboratory at Dartmouth within 24 h for processing and analysis.

Genotyping

DNA was isolated using Qiagen genomic DNA extraction kits (QIAGEN Inc, Valencia, CA) from peripheral circulating blood lymphocyte specimens harvested at the time of interview. Genotyping was performed on all DNA samples of sufficient concentration (893 controls, 617 cases), using the Cancer Panel on the GoldenGate Assay system by Illumina's Custom Genetic Analysis service (Illumina Inc., San Diego, CA). The Cancer Panel contains 1,421 SNPs in approximately 400 hypothesized cancer-related genes from the SNP500 database (<http://www.illumina.com>). Genotype data were available on 75% of interviewed subjects and the genotyped population had similar characteristics as the overall study. Samples repeated on multiple GoldenGate plates yielded the same call for 99.9% of SNPs and 99.5% of samples submitted were successfully genotyped. Genotype calls were 99% concordant between genotyping platforms.

Statistical analysis

The main goals of the statistical analysis were to identify SNPs and combinations of SNPs that relate to bladder cancer susceptibility or survival. Selection of SNPs with dose-responsive relationships to bladder cancer risk was performed by ranking SNPs based on *p*-for-trend values associated with a dose response of variant alleles with FDR value ≤ 0.56 .

Kullback–Leibler distance measures were used to compare discrete distributions of feature values for cases and controls (Duda et al. 2001). The Kullback–Leibler divergence provides an information-based measure of the difference between two probability distributions. The Kullback–Leibler divergence may be expressed as the number of bits of information required to encode samples based on the two distributions. The Kullback–Leibler divergence differs from a conventional distance metric on the space of probability distributions in that it is not symmetric and does not satisfy the triangle inequality. As an information measure, the Kullback–Leibler divergence may be thought of as a non-linear measure that often accentuates distribution difference properties that are not detected by conventional distance metrics (Kullback 1959, 1987). While the divergence cannot be easily generalized in its performance, it is often a valuable complement to conventional distance metrics in applications where well-defined models are not easily available. Kullback–Leibler divergence has recently been to be used in studies of genetic association [e.g. (Cussenot et al. 2008)]. The SNPs were ranked and the top 20 SNPs with the largest scores were selected.

Two, three, and four-way combinations of *p*-for-trend and Kullback–Leibler selected potentially interacting SNPs were identified using Multifactor Dimensionality Reduction (MDR) analysis (Ritchie et al. 2001). To allow for epistatic interactions in the absence of main effects, we used a more liberal *p*-for-trend cutoff for entry into the MDR model (FDR value 0.83). MDR is a data reduction (i.e. constructive induction) approach that seeks to identify combinations of multilocus genotypes and discrete environmental factors that are associated with either high risk or low risk of disease. Thus, MDR defines a single variable that

incorporates information from several loci and/or environmental factors that can be divided into high risk and low risk combinations. This new variable can be evaluated for its ability to classify and predict outcome risk status using cross-validation and permutation testing (Ritchie et al. 2001, 2003).

We also used the tuned Relief F as a pre-filtering step to help identify epistatic or non-additive, interacting SNPs in MDR. ReliefF considers SNPs jointly rather than individually. In our experience, this method is better for detecting interactions than chi-square test. ReliefF estimates weights for each SNP “estimated based on whether the nearest neighbors (nearest hit, *H*) of a randomly selected instance from the same class and the nearest neighbors from the other class (nearest miss, *M*) have the same or different values” (Moore and White 2007). Tuned ReliefF uses vectors of SNPs, removes those of low quality, re-estimates, and then ranks them based on their ability to discriminate case-control status.

To assess the magnitude of the independent effects of the SNPs that were identified by the filters described above, we conducted logistic regression analyses for individuals with one or two variant alleles in comparison to those homozygous wild type for each individual SNP. These analyses were adjusted for age (less than or greater than median age 64), gender, and smoking status (never, former, current). We evaluated the nature of interactions between bladder cancer risk factors (gender, smoking, age) and/or the SNP allele combinations predicted by MDR, by including interaction terms in a logistic regression model. Statistical significances of the interactions were assessed using likelihood ratio tests comparing the models with and without the interaction terms (e.g. $\text{logit}(p) = \text{constant} + \text{SNP1} + \text{SNP2}$ vs. $\text{logit}(p) = \text{constant} + \text{SNP1} + \text{SNP2} + \text{interaction}$). MDR identifies combinations of predictive factors, but does not distinguish between multiplicative and additive effects. Thus, we used the likelihood ratio test to assess the multiplicative nature of the interactions. *p* values represent two-sided statistical tests with statistical significance at $P < 0.05$. Our analytic strategy embraces both the algorithmic and traditional statistical approaches [discussed in (Breiman 2001)], allowing us to identify epistatic SNP interactions.

Selection of SNPs related to bladder cancer survival was performed using an FDR threshold < 0.5 , choosing the four individual SNPs with the lowest log-rank *p* values (≤ 0.1). Then we used Gene Set Assessment (GSA) to investigate groups of SNPs that are members of a common gene. The GSA method calculates an enrichment score based on whether the gene-based SNP sets contain a number of individual SNPs that are highly associated with cancer (over-represented), with adjustment for multiple testing (Holden et al. 2008). GSA was used to select the top 10 ranking genes (each containing multiple SNPs) by log-rank *p* values. Survival plots for bladder cancer cases were generated using the Kaplan-Meier method and differences between genotypes were assessed using the log-rank test. To adjust for additional factors related to patient survival, Cox-proportional hazards regression analysis was performed with age, gender, smoking (never, former, current), as well as AJCC tumor stage (0, 0is, 1–4), grade (1–4), size ($<$, ≥ 30 mm), treatment (immunotherapy with surgery; chemotherapy, radiotherapy, or immunotherapy only; chemotherapy or radiotherapy with cystectomy; transurethral resection only; cystectomy only) in the model. Results were not sensitive to removal of subjects with metastatic tumors from the analysis (data not shown). *p* values represent two-sided statistical tests with statistical significance at $P < 0.05$.

Results

The characteristics of the subjects who provided a DNA sample were similar to those of the overall study population. The majority of the population was Caucasian. The case group contained a higher percentage of current smokers than the controls (Table 1). A large proportion of the cases were non-invasive bladder cancer, reflecting the population-based nature of the

study. Testing showed deviation from Hardy–Weinberg equilibrium for ADH1C_14, and MSH6_01.

Risk models

The *p*-for-trend method that hypothesizes a dose–response effect by the number of variant alleles was used to select the individual SNPs shown in Table 2. Individuals homozygous for the variant forms of the alcohol dehydrogenase SNP ADH1C_14 had an increased risk of bladder cancer compared with wild type [OR 1.7 95% CI (1.3–2.2)]. Increased risk was also observed among individuals with the homozygous variant form of SNPs in *MTHFD2*, *TEP1*, *AURKA*, *FZD7*, *IGF1*, *MSH6*. Gene variants associated with reduced risk included *IL8RB*, *MET*, *SLC4A2*, *MBL2* (Table 2). Of the SNPs selected in the *p*-for-trend analysis, only the risk associated with ADH1C_14 varied by smoking status. Never smokers had modestly higher bladder cancer risk OR 1.4 (95% CI 0.9–2.3) than current smokers OR 0.9 (95% CI 0.6–1.4) (gene–current smoker interaction *p* = 0.05). We also analyzed these SNPs separated by study phase (I, II). Results within each phase were consistent with our analysis of the overall population.

To identify SNPs that do not follow the *p*-for-trend risk model, we used the Kullback–Leibler distance algorithm to rank SNPs for an association with bladder cancer risk (heterozygous (het.) or variant (var.) vs. wildtype). Of the top 20 SNPs selected by Kullback–Leibler distance, five showed individual gene relationships with bladder cancer incidence when assessed by logistic regression analysis after adjustment for age, gender, and smoking status, GSTZ1_02, AKR1C3_35, TYR_02, SCARB1_03, and SLC23A1_05 (Table 3). The other SNPs that had top ranking Kullback–Leibler distances did not show a significant association with bladder cancer by logistic regression analysis with adjustment for potential confounders, including SCARB1_03, PLA2G6_08, PTGS1_02, EPHX1_17, SLC23A2_31, SLC6A3_14, TGM1_01, AKR1C3_11, KRAS_22, CTNNB1_02, IL10_06, HMGCR_01, BARD1_02, AXIN2_09, PTGS1_02, MSH2_06 (data not shown).

We then identified potential interactions between these selected SNPs using MDR, followed by quantification of the odds ratios associated with the high versus low risk combinations of alleles/characteristics using logistic regression analysis with adjustment for age, gender and smoking status. Of the *p*-for-trend selected genes using an FDR value cutoff of 0.8, MDR indicated that a combination of the high-risk genotypes for PIN1_21 rs889162 g.IVS3+2592T>C, MET_01 rs41736 g.Ex20+60C>T p.D1304D, and AMACR_17 rs10941112 g.Ex3-29A>G p.G175D was significantly associated with bladder cancer (MDR accuracy 0.56, cross-validation consistency 7/10). The logistic regression adjusted OR was 2.3 [95% CI 1.8–2.9] for the high risk alleles, versus the low risk genotype combination. Of these PIN1_21 and AMACR_17 indicated the strongest evidence of synergy (likelihood ratio test gene–gene interaction *p* value = 0.02).

We also entered the top ranking 20 Kullback–Leibler distance selected genes, along with age, gender and smoking status, into MDR. The combination of high-risk genotypes for SLC23A2_31 rs12479919 g.IVS1+1312G>A, AKR1C3_11 rs2275928 g.IVS8+40A>G, SCARB1_03 rs4765621 g.IVS1-18462G>A, and PLA2G6_08 rs2016755 g.IVS3-309T>C was associated with elevated bladder cancer risk (logistic regression adjusted OR 2.6 (95% CI 2.1–3.2), MDR accuracy 0.53, cross-validation consistency 6/10). Within this MDR-selected group, we tested possible combinations of gene–gene interactions and found a significant relationship between SCARB1_03 and AKR1C3_11 (interaction *p* value = 0.04). This MDR analysis also picked the strongest two-factor interaction from the Kullback–Leibler selected genes, between AKR1C3_11 and IL10_06 (MDR accuracy 0.56, cross-validation consistency 6/10), adjusted OR 1.6 (95% CI 1.3–2.0), gene–gene interaction *p* value = 0.0004. Using the tuned ReliefF filter in MDR, we also identified CTNNB1_02 rs11564452 g.IVS7-562A>T,

SCARB1_03 rs4765621 g.IVS1-18462G>A, SLC23A2_31 rs12479919 g.IVS1+1312G>A, and BARD1_02 rs1129804 g.Ex1+44C>G as a high risk combination (MDR accuracy 0.56, cross-validation consistency 10/10). The logistic regression adjusted OR was 2.3 (95% CI 1.7–2.9) for the high versus low risk allele combination. Of these, SCARB1_03 and SLC23A2_31 showed the strongest evidence of a gene–gene effect (interaction $P = 0.02$) on bladder cancer risk.

Additionally, we analyzed the data using a χ^2 filter on the entire dataset. Entering the χ^2 top-ranked 40 genes into MDR revealed an interaction between SLC19A1_01 rs1051266 g.Ex4-114T>C p.H27R, and IGF2AS_04 rs3741212 g.Ex1+112A>G, MDR accuracy 0.60, cross-validation consistency 10/10. The combined any variant adjusted odds ratio for this gene–gene combination from logistic regression was 1.4 (95% CI 1.1–1.7) (interaction $P < 0.0001$).

Survival models

We used the log-rank test to identify the top four individual SNPs likely to be related to bladder cancer case survival (Table 4). Cox-regression analysis (adjusted for age, gender, smoking status, stage/grade, and treatment) suggested that lower survival was associated with having SNPs in the T lymphocyte regulator CD80_04, the apoptosis regulator BCL2L1_03, and the ERCC4_01 nucleotide excision repair gene. The polycyclic aromatic hydrocarbon (PAH) metabolism SNP, EPHX1_15 was associated with better prognosis (Fig. 1). The effects did not vary significantly by smoking status (e.g. ERCC4 among never smokers HR 10.6 (95% CI 2.8–38), ever smokers HR 1.9 (95% CI 1.2–3.1), interaction $P = 0.2$).

Gene Set Assessment allowed us to use all the SNP data from each gene as a group to identify the genes that most strongly predicted survival by log-rank test (Table 5). We then evaluated the individual SNPs within the top 10 ranked genes using adjusted Cox-regression models. The variant form of the transcriptional regulator GATA3_29 was associated with poor survival HR 2.8 (95% CI 1.6–5.0). The same apoptotic gene BCL2L1 selected by this GSA analysis was also picked in the individual SNP log-rank test (BCL2L1_03). An additional SNP in this gene, BCL2L1_01, was also related to poor bladder cancer survival [any variant HR 1.6 (1.1–2.2)]. This GSA analysis also predicted poor survival for individuals with SNPs in the angiogenic gene IL8RB_01 [any variant HR 1.9 (1.3–2.7)] (Fig. 1), the mitotic genes PLK1_15 [any variant HR 1.3 (0.9–1.8)] and STK6_06 [any variant HR 1.4 (0.9–1.9)], the metabolism gene UGT1A1_24 [any variant HR 1.4 (1.0–1.9)], the apoptotic gene CASP9_01 [any variant HR 1.7 (1.2–2.4)] (Fig. 1) and better survival with the steroid receptor PGR_05 [any variant HR 0.7 (0.5–1.0)] (Fig. 1), and the telomerase TERT_02 [any variant HR 0.8 (0.6–1.2)]. The selected SNPs did not differ significantly by smoking status (data not shown).

Discussion

A genetic component for bladder cancer susceptibility is likely since individuals with a family history have a higher incidence. Furthermore, although cigarette smoking is the strongest risk factor, it is not clear why some smokers get cancer and others do not (Ahsan and Thomas 2004). The magnitude of the association between the SNPs that are well accepted to be predictive (such as the N-acetyltransferases and DNA repair genes) and bladder cancer risk is relatively small. Thus, there are likely to be other causal genetic factors that remain to be discovered. Bladder cancer survival also varies widely between individual patients, even within the same stage and grade of the tumor. We hypothesize that this survival is in part genetically controlled, yet reliable predictors have not been identified or validated.

The occurrence and prognosis for bladder cancer and other relatively common diseases is most likely controlled by a combination of multiple genetic factors and exposures, rather than a single polymorphism. Our strategy was to predict the single SNPs that have the strongest

relationships with bladder cancer incidence or survival. The magnitude of these relationships was then evaluated using traditional logistic regression analysis procedures. We then identified the combinations of the top ranked SNPs and genes using multi-factor selection algorithms for the most predictive high risk sets of alleles. These models were then assessed statistically by traditional methods. Using this approach, we identified a number of SNPs and SNP combinations that were associated with bladder cancer incidence and survival. Modifying SNPs occurred in biologically plausible genes that regulate metabolism, DNA repair, telomere maintenance, mitosis, inflammation, and apoptosis.

Alcohol dehydrogenase 1C (ADH1C) metabolizes multiple substrates, most notably ethanol. Consistent with our data, a previous study in the Netherlands suggested an increased risk of bladder cancer among variants, but no interaction with alcohol consumption (van Dijk et al. 2001). UGT1A1 is a phase 2 enzyme which metabolizes lipophilic molecules (e.g. steroids, bilirubin, hormones, drugs, and aromatic amines such as benzidine) via glucuronidation for excretion (Zenser et al. 2002) and was related to worse survival. Polymorphisms in the one-carbon methyl-metabolism gene methylenetetrahydrofolate reductase (*MTHFR*) have been well studied in relation to bladder cancer risk (Karagas et al. 2005; Moore et al. 2004; Ouerhani et al. 2007; Vineis et al. 2007). The methyl-group metabolism pathway regulates methylation of DNA CpG islands, thus controlling expression and silencing of tumor suppressor and oncogenes (Paz et al. 2002). Our study found increased risk for variants in a related methyl-metabolism dehydrogenase enzyme that provides the 5,10 methylenetetrahydrofo-late pool (*MTHFD2*). This result is consistent with studies showing that *MTHFD2* SNPs modify gastric and lung cancer risk (Liu et al. 2008; Wang et al. 2007).

Telomerase activity is critical for chromosomal stability and has been associated with bladder cancer risk (Alvarez and Lokeshwar 2007; Xing et al. 2007). TERT1 protein levels are a rate limiting step in telomerase activity and were associated with better survival in a previous study of 132 bladder cancer cases ($P = 0.007$) (Ito et al. 1998; Mavrommatis et al. 2005). Our study indicated better survival for *TERT* variants. Variants of telomerase-associated protein 1 (*TEPI*), a telomerase that adds new telomeres to the ends of chromosomes, protecting them from degradation, were associated with increased bladder cancer risk in our study, and in other cancers in a prior study (Savage et al. 2007).

Improper distribution of chromosomes during mitosis can result from modified expression of mitotic kinases, including Aurora kinase A (*AURKA*), and leads to chromosomal instability and aneuploidy in bladder cells (Fraizer et al. 2004). Subjects with the *AURKA* non-synonymous polymorphism in exon 11 had significantly greater susceptibility to bladder cancer in our study. Allelic variation in *AURKA* has been associated with *AURKA* expression levels and increased risk of multiple cancers in a meta-analysis of 9,549 cases (Ewart-Toland et al. 2005; Matarasso et al. 2007). *AURKA* amplification is very prevalent in bladder tumors (64%) and tumor *AURKA* gene and protein expression levels are associated with tumor recurrence and shorter survival time (Comperat et al. 2007; Denzinger et al. 2007; Schultz et al. 2007). The previously examined *AURKA* polymorphism T91A was not associated with bladder cancer prognosis in a study of 135 patients (Schultz et al. 2007). Our results suggest a lower survival rate for variants at another locus in *AURKA* intron 7 and a possible relationship with polo-like kinase 1 (*PLK1*). Aurora kinase A activates *PLK1*, which regulates spindle formation as DNA damage arrested cells move into mitosis (Macurek et al. 2008). High *PLK1* expression levels are associated with bladder tumor chromosomal instability and progression (Yamamoto et al. 2006).

Insulin-like growth factor 1 (*IGF1*) variants had increased risk of bladder cancer in our study. SNPs in the promoter and intronic regions of *IGF1* predict circulating IGF1 levels and increase colorectal cancer risk (Palles et al. 2008; Wong et al. 2008). Insulin-like growth factor 2

antisense (*IGF2AS/PEG1*) is aberrantly methylated in colorectal cancers and is overexpressed in Wilms' tumors (Nishihara et al. 2000; Okutsu et al. 2000). *IGF2AS* variants interacted with the folate transporter, *SLC19A1*, to increase bladder cancer risk in our study. Similarly, *SLC19A1* was related to increased bladder cancer risk in the Spanish population (Moore et al. 2007).

Bladder cancer survival was negatively associated with variation in *CD80*, one of the surface antigens that increase in expression following treatment with the immunotherapy drug bacillus Calmette-Guerin (BCG) that is commonly used to treat bladder cancer patients (Ikeda et al. 2002). Interleukin-8 receptor (*IL8RB/CXCR2*) variants were at decreased risk of developing bladder cancer, but also experienced poorer survival than wildtype individuals in our study. *IL8RB* is associated with inflammatory response mediating angiogenesis, and neutrophil migration (Matheson et al. 2006). *IL8RB* polymorphisms have previously been associated with breast cancer progression (Kamali-Sarvestani et al. 2007).

Our data also suggest an increased risk of bladder cancer among variants for the mutS homolog 6 (*MSH6*), which recognizes mismatched nucleotides and induces apoptosis if repair fails. Reduced expression, mutation of *MSH6* and concomitant loss of mismatch repair activity have been observed in bladder tumors (Mongiat-Artus et al. 2006; Thykjaer et al. 2001). As in our study, *MSH6* variants were at higher risk for malignant bladder, an observation made for other cancers as well (Campbell et al. 2008; Sanyal et al. 2007). We observed shorter survival among *ERCC4* variants. Although genetic variation in the DNA repair gene *ERCC4* was not independently associated with bladder cancer susceptibility, few have investigated its role in survival (Garcia-Closas et al. 2006; Matullo et al. 2005).

Variants in apoptosis genes also modulated bladder cancer survival. The BCL-2 family member, BCL2L1/BCLX regulates the mitochondrial membrane potential controlling release of apoptotic inducers and expression levels have been related to cancer progression (Gazzaniga et al. 1998; Kirsh et al. 1998). Variants of the early apoptosis cascade member cysteine peptidase, Caspase 9 (*CASP9*) experienced shorter survival in our study. Both BCL2L1 and caspase 9 may be involved in response to some types of chemotherapy treatment in the bladder (Yuan et al. 2002).

In addition to main-gene eVects, we also detected a possible interaction between peptidylprolyl cis/trans isomerase (*PIN1*) and Alpha-methylacyl-CoA racemase (*AMACR*) variants. *PIN1* is a conditional tumor suppressor, which transduces phosphorylation of P53 and promotes cell death by mediating dissociation from the apoptosis inhibitor iASPP (Mantovani et al. 2007). *PIN1* plays a critical role in the DNA damage checkpoint, in part by protecting p53 from the inhibitor MDM2. Thus, *Pin1*^{-/-} cells do not undergo mitosis in response to DNA damage (Takahashi et al. 2008). *AMACR* is involved in cellular energy metabolism by the oxidation of branched-chain fatty acids and in the metabolism of lipophilic drugs (Lloyd et al. 2008). *AMACR* expression is associated with higher bladder tumor histopathologic grade, implying that endogenous fatty acids may be used for energy as tumors grow (Gunia et al. 2008). *AMACR* variations have previously been associated with prostate cancer risk (Levin et al. 2007). The combination of inhibition of the DNA damage checkpoint by *PIN1* with *AMACR* modulated energy availability for tumor growth and could result in the synergistic effect on tumor risk that we observed.

Another interaction involved scavenger receptor class B (*SCARB1*), which regulates cellular cholesterol uptake and sequesters cholesteryl esters from high-density lipoproteins (Miquel et al. 2003) with the aldo-keto reductase, *AKR1C3*, which is strongly expressed in the bladder. *AKR1C3* regulates lipophilic ligand access to hormone receptors and activates PAHs, which are known bladder carcinogens (Azzarello et al. 2008). Our finding of modified bladder cancer

risk in New Hampshire corroborates the recently reported association between *AKR1C3* polymorphisms and bladder cancer risk in Spain (Figuerola et al. 2008).

The applicability of these results to other populations of bladder cancer patients needs to be explored before these markers can be used to guide clinical practice. These data will be used to inspire investigation of the mechanistic impact of these SNPs on tumor biology. Intronic SNPs may be acting through alternative splice sites, miRNAs, conformational changes or linkage disequilibrium with another functional SNP. Based on our results, a thorough investigation into SNPs in the metabolism, telomere maintenance, mitosis, inflammation, and apoptosis pathways is warranted. Nevertheless, these findings highlight a number of genetic variations in known carcinogenic pathway genes that modify bladder cancer risk and prognosis. These results may provide new insight into methods of predicting bladder cancer susceptibility using SNP panels, identify targets for bladder cancer prevention, and guide individualized choice of treatment methodologies based on molecular pathway alterations.

Acknowledgments

This publication was funded in part by grant numbers CA102327, CA121382, CA099500, CA82354, CA57494, CA078609, ES00002, P42 ES05947, RR018787, LM009012 and ES007373 from the National Cancer Institute, NIH and from the National Institute of Environmental Health Sciences, NIH.

References

- Ahsan H, Thomas DC. Lung cancer etiology: independent and joint effects of genetics, tobacco, and arsenic. *JAMA* 2004;292:3026–3029. [PubMed: 15613673]
- Alvarez A, Lokeshwar VB. Bladder cancer biomarkers: current developments and future implementation. *Curr Opin Urol* 2007;17:341–346. [PubMed: 17762628]
- Andrew AS, Nelson HH, Kelsey KT, et al. Concordance of multiple analytical approaches demonstrates a complex relationship between DNA repair gene SNPs, smoking and bladder cancer susceptibility. *Carcinogenesis* 2006;27:1030–1037. [PubMed: 16311243]
- Azzarello J, Fung KM, Lin HK. Tissue distribution of human *AKR1C3* and rat homolog in adult genitourinary system. *J Histochem Cytochem* 2008;56(9):853–861. [PubMed: 18574251]
- Breiman L. Statistical modeling the two cultures. *Stat Sci* 2001;16:199–231.
- Campbell PT, Curtin K, Ulrich C, et al. Mismatch repair polymorphisms and risk of colon cancer, tumor microsatellite instability, and interactions with lifestyle factors. *Gut*. 2008doi:10.1136/gut.2007.144220
- Carroll P, Raghavan D, Stein J, et al. H0005 PM-the treatment of bladder cancer. Stage by Stage. AUA Annual Meeting 2001. 2001
- Comperat E, Camparo P, Haus R, et al. Aurora-A/STK-15 is a predictive factor for recurrent behaviour in non-invasive bladder carcinoma: a study of 128 cases of non-invasive neoplasms. *Virchows Arch* 2007;450:419–424. [PubMed: 17333265]
- Cussenot O, Azzouzi AR, Bantsimba-Malanda G, et al. Effect of genetic variability within 8q24 on aggressiveness patterns at diagnosis and familial status of prostate cancer. *Clin Cancer Res* 2008;14:5635–5639. [PubMed: 18765558]
- Denzinger S, Stoehr R, Schwarz S, et al. Low level STK15 amplification in histologically benign urothelium of patients with bladder cancer adversely predicts patient outcome following cystectomy. *Int J Oncol* 2007;31:793–802. [PubMed: 17786310]
- Duda, RO.; Hart, PE.; Stork, DG. Pattern classification. Vol. 2nd edn. Wiley; New York: 2001. p. 632
- Ewart-Toland A, Dai Q, Gao YT, et al. Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis* 2005;26:1368–1373. [PubMed: 15802297]
- Figuerola JD, Malats N, Garcia-Closas M, et al. Bladder cancer risk and genetic variation in *AKR1C3* and other metabolizing genes. *Carcinogenesis* 2008;29:1955–1962. [PubMed: 18632753]

- Fraizer GC, Diaz MF, Lee IL, et al. Aurora-A/STK15/BTAK enhances chromosomal instability in bladder cancer cells. *Int J Oncol* 2004;25:1631–1639. [PubMed: 15547700]
- Garcia-Closas M, Malats N, Silverman D, et al. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 2005;366:649–659. [PubMed: 16112301]
- Garcia-Closas M, Malats N, Real FX, et al. Genetic variation in the nucleotide excision repair pathway and bladder cancer risk. *Cancer Epidemiol Biomarkers Prev* 2006;15:536–542. [PubMed: 16537713]
- Gazzaniga P, Gradilone A, Silvestri I, et al. Variable levels of bcl-2, bcl-x and bax mRNA in bladder cancer progression. *Oncol Rep* 1998;5:901–904. [PubMed: 9625842]
- Gunia S, May M, Scholmann K, et al. Expression of alpha-methylacyl-CoA racemase correlates with histopathologic grading in noninvasive bladder cancer. *Virchows Arch* 2008;453:165–170. [PubMed: 18648853]
- Holden M, Deng S, Wojnowski L, et al. GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies. *Bioinformatics* 2008;24:2784–2785. [PubMed: 18854360]
- Ikeda N, Toida I, Iwasaki A, et al. Surface antigen expression on bladder tumor cells induced by bacillus Calmette-Guerin (BCG): a role of BCG internalization into tumor cells. *Int J Urol* 2002;9:29–35. [PubMed: 11972647]
- Ito H, Kyo S, Kanaya T, et al. Expression of human telomerase subunits and correlation with telomerase activity in urothelial cancer. *Clin Cancer Res* 1998;4:1603–1608. [PubMed: 9676833]
- Jemal A, Siegel R, Ward E, et al. Cancer statistics. *CA Cancer J Clin* 2007;57:43–66. [PubMed: 17237035]
- Kamali-Sarvestani E, Aliparasti MR, Atefi S. Association of interleukin-8 (IL-8 or CXCL8) -251T/A and CXCR2 +1208C/T gene polymorphisms with breast cancer. *Neoplasma* 2007;54:484–489. [PubMed: 17949231]
- Kantor AF, Hartge P, Hoover RN, et al. Familial and environmental interactions in bladder cancer risk. *Int J Cancer* 1985;35:703–706. [PubMed: 4008097]
- Karagas MR, Tosteson TD, Blum J, et al. Design of an epidemiologic study of drinking water arsenic exposure and skin and bladder cancer risk in a U.S. population. *Environ Health Perspect* 1998;106 (Suppl 4):1047–1050. [PubMed: 9703491]
- Karagas MR, Park S, Nelson HH, et al. Methylenetetrahydrofolate reductase (MTHFR) variants and bladder cancer: a population-based case-control study. *Int J Hyg Environ Health* 2005;208:321–327. [PubMed: 16217917]
- Kirkali Z, Chan T, Manoharan M, et al. Bladder cancer: epidemiology, staging and grading, and diagnosis. *Urology* 2005;66:4–34. [PubMed: 16399414]
- Kirsh EJ, Baunoch DA, Stadler WM. Expression of bcl-2 and bcl-X in bladder cancer. *J Urol* 1998;159:1348–1353. [PubMed: 9507882]
- Kullback, S. *Information theory and statistics*. Wiley; New York: 1959.
- Kullback S. The Kullback-Leibler distance. *Am Stat* 1987;41:340–341.
- Levin AM, Zuhlke KA, Ray AM, et al. Sequence variation in alpha-methylacyl-CoA racemase and risk of early-onset and familial prostate cancer. *Prostate* 2007;67:1507–1513. [PubMed: 17683075]
- Liu H, Jin G, Wang H, et al. Association of polymorphisms in one-carbon metabolizing genes and lung cancer risk: a case-control study in Chinese population. *Lung Cancer* 2008;61:21–29. [PubMed: 18221821]
- Lloyd MD, Darley DJ, Wierzbicki AS, et al. Alpha-methylacyl-CoA racemase—an ‘obscure’ metabolic enzyme takes centre stage. *FEBS J* 2008;275:1089–1102. [PubMed: 18279392]
- Macurek L, Lindqvist A, Lim D, et al. Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature* 2008;455:119–123. [PubMed: 18615013]
- Mantovani F, Tocco F, Girardini J, et al. The prolyl isomerase Pin1 orchestrates p53 acetylation and dissociation from the apoptosis inhibitor iASPP. *Nat Struct Mol Biol* 2007;14:912–920. [PubMed: 17906639]
- Matarasso N, Bar-Shira A, Rozovski U, et al. Functional analysis of the Aurora Kinase A Ile31 allelic variant in human prostate. *Neoplasia* 2007;9:707–715. [PubMed: 17898866]

- Matheson MC, Ellis JA, Raven J, et al. Association of IL8, CXCR2 and TNF-alpha polymorphisms and airway disease. *J Hum Genet* 2006;51:196–203. [PubMed: 16429233]
- Matullo G, Guarrera S, Sacerdote C, et al. Polymorphisms/Haplotypes in DNA repair genes and smoking: a bladder cancer case-control study. *Cancer Epidemiol Biomarkers Prev* 2005;14:2569–2578. [PubMed: 16284380]
- Mavrommatis J, Mylona E, Gakiopoulou H, et al. Nuclear hTERT immunohistochemical expression is associated with survival of patients with urothelial bladder cancer. *Anticancer Res* 2005;25:3109–3116. [PubMed: 16080574]
- Miquel JF, Moreno M, Amigo L, et al. Expression and regulation of scavenger receptor class B type I (SR-BI) in gall bladder epithelium. *Gut* 2003;52:1017–1024. [PubMed: 12801960]
- Mongiart-Artus P, Miquel C, Van der Aa M, et al. Microsatellite instability and mutation analysis of candidate genes in urothelial cell carcinomas of upper urinary tract. *Oncogene* 2006;25:2113–2118. [PubMed: 16288216]
- Moore JH, White BC. Tuning ReliefF for Genome-Wide Genetic Analysis. *Lect Notes Comput Sci* 2007;4447:166–175.
- Moore LE, Wiencke JK, Bates MN, et al. Investigation of genetic polymorphisms and smoking in a bladder cancer case-control study in Argentina. *Cancer Lett* 2004;211:199–207. [PubMed: 15219943]
- Moore LE, Malats N, Rothman N, et al. Polymorphisms in one-carbon metabolism and trans-sulfuration pathway genes and susceptibility to bladder cancer. *Int J Cancer* 2007;120:2452–2458. [PubMed: 17311259]
- Murta-Nascimento C, Schmitz-Drager BJ, Zeegers MP, et al. Epidemiology of urinary bladder cancer: from tumor development to patient's death. *World J Urol* 2007a;25:285–295. [PubMed: 17530260]
- Murta-Nascimento C, Silverman DT, Kogevinas M, et al. Risk of bladder cancer associated with family history of cancer: do low-penetrance polymorphisms account for the increase in risk? *Cancer Epidemiol Biomarkers Prev* 2007b;16:1595–1600. [PubMed: 17684133]
- Nishihara S, Hayashida T, Mitsuya K, et al. Multipoint imprinting analysis in sporadic colorectal cancers with and without microsatellite instability. *Int J Oncol* 2000;17:317–322. [PubMed: 10891541]
- Okutsu T, Kuroiwa Y, Kagitani F, et al. Expression and imprinting status of human PEG8/IGF2AS, a paternally expressed antisense transcript from the IGF2 locus, in Wilms' tumors. *J Biochem* 2000;127:475–483. [PubMed: 10731720]
- Ouerhani S, Oliveira E, Marrakchi R, et al. Methylenetetrahydrofolate reductase and methionine synthase polymorphisms and risk of bladder cancer in a Tunisian population. *Cancer Genet Cytogenet* 2007;176:48–53. [PubMed: 17574963]
- Palles C, Johnson N, Coupland B, et al. Identification of genetic variants that influence circulating IGF1 levels: a targeted search strategy. *Hum Mol Genet* 2008;17:1457–1464. [PubMed: 18250100]
- Paz MF, Avila S, Fraga MF, et al. Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors. *Cancer Res* 2002;62:4519–4524. [PubMed: 12154064]
- Ritchie MD, Hahn LW, Roodi N, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 2001;69:138–147. [PubMed: 11404819]
- Ritchie MD, Hahn LW, Moore JH. Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. *Genet Epidemiol* 2003;24:150–157. [PubMed: 12548676]
- Sanyal S, De Verdier PJ, Steineck G, et al. Polymorphisms in XPD, XPC and the risk of death in patients with urinary bladder neoplasms. *Acta Oncol* 2007;46:31–41. [PubMed: 17438703]
- Savage SA, Chanock SJ, Lissowska J, et al. Genetic variation in five genes important in telomere biology and risk for breast cancer. *Br J Cancer* 2007;97:832–836. [PubMed: 17848914]
- Schned AR, Andrew AS, Marsit CJ, et al. Survival following the diagnosis of noninvasive bladder cancer: WHO/International Society of Urological Pathology Versus WHO Classification Systems. *J Urol* 2007;178:1196–1200. [PubMed: 17698105]

- Schultz IJ, Kiemeny LA, Roelofs R, et al. The prognostic role of the STK15 T91A polymorphism and of STK15 mRNA expression in patients with urothelial cell carcinoma. *Anticancer Res* 2007;27:1025–1030. [PubMed: 17465238]
- Takahashi K, Uchida C, Shin RW, et al. Prolyl isomerase, Pin1: new findings of post-translational modifications and physiological substrates in cancer, asthma and Alzheimer's disease. *Cell Mol Life Sci* 2008;65:359–375. [PubMed: 17965833]
- Thykjaer T, Christensen M, Clark AB, et al. Functional analysis of the mismatch repair system in bladder cancer. *Br J Cancer* 2001;85:568–575. [PubMed: 11506498]
- van Dijk B, van Houwelingen KP, Witjes JA, et al. Alcohol dehydrogenase type 3 (ADH3) and the risk of bladder cancer. *Eur Urol* 2001;40:509–514. [PubMed: 11752857]
- Vineis P, Veglia F, Garte S, et al. Genetic susceptibility according to three metabolic pathways in cancers of the lung and bladder and in myeloid leukemias in nonsmokers. *Ann Oncol* 2007;18:1230–1242. [PubMed: 17496311]
- Wang L, Ke Q, Chen W, et al. Polymorphisms of MTHFD, plasma homocysteine levels, and risk of gastric cancer in a high-risk Chinese population. *Clin Cancer Res* 2007;13:2526–2532. [PubMed: 17438114]
- Wong HL, Koh WP, Probst-Hensch NM, et al. Insulin-like growth factor-1 promoter polymorphisms and colorectal cancer: a functional genomics approach. *Gut* 2008;57:1090–1096. [PubMed: 18308828]
- Wu X, Gu J, Grossman HB, et al. Bladder cancer predisposition: a multigenic approach to DNA-repair and cell-cycle-control genes. *Am J Hum Genet* 2006;78:464–479. [PubMed: 16465622]
- Xing J, Zhu Y, Zhao H, et al. Differential induction in telomerase activity among bladder cancer patients and controls on gamma-radiation. *Cancer Epidemiol Biomarkers Prev* 2007;16:606–609. [PubMed: 17372259]
- Yamamoto Y, Matsuyama H, Kawauchi S, et al. Overexpression of Polo-Like Kinase 1 (PLK1) and chromosomal instability in bladder cancer. *Oncology* 2006;70:231–237. [PubMed: 16837776]
- Yuan SY, Hsu SL, Tsai KJ, et al. Involvement of mitochondrial pathway in Taxol-induced apoptosis of human T24 bladder cancer cells. *Urol Res* 2002;30:282–288. [PubMed: 12389115]
- Zenser TV, Lakshmi VM, Hsu FF, et al. Metabolism of N-acetylbenzidine and initiation of bladder cancer. *Mutat Res* 2002;506–507:29–40.

Fig. 1. Bladder cancer survival in relation to SNPs. Kaplan–Meier plots show survival (y-axis) versus years from diagnosis (x-axis) by **a** EPHX1_15 genotype, **b** IL8RB_01 genotype, **c** PGR_05 genotype, and **d** CASP9_01 genotype. *Black lines* are homozygous wildtype, *red* are homozygous variant, *green* are heterozygous

Table 1
Genotyped population characteristics

	Controls <i>n</i> (%)	Cases <i>n</i> (%)	<i>p</i> value
Gender			
Female	338 (38)	150 (24)	Ref.
Male	555 (62)	467 (76)	<0.0001
Reference age			
<40	43 (5)	17 (3)	Ref.
40–55	152 (17)	102 (17)	0.4
55–70	480 (54)	337 (54)	0.3
>70	218 (24)	161 (26)	0.2
Race			
White	889 (99)	607 (98)	Ref.
Non-white	4 (1)	10 (2)	0.06
Smoking status			
Never	305 (34)	106 (17)	Ref.
Former	447 (50)	305 (50)	<0.0001
Current	141 (16)	205 (33)	<0.0001
AJCC Stage			
0	–	337 (56)	
0is	–	24 (4)	
1	–	170 (28)	
2	–	43 (7)	
3, 4	–	33 (5)	

Data were missing for smoking on 1 case, stage on 10 cases

Table 2
Individual SNPs selected for association with bladder cancer risk by *p*-for-trend <0.004

SNP (function)	Genotype	Controls		Cases		N	Bladder cancer OR (95% CI)	
		N	%	N	%		N	%
ADH1C_14 (Alcohol dehydrogenase)	wt	440	51	256	431.0 Ref.			
rs2241894	het	264	31	181	311.2 (0.9–1.5)			
g-Ex5+106A>G	var	156	18	152	261.7 (1.3–2.2)			
p-T151T	any var	420	49	333	571.4 (1.1–1.7)			
MTHFD2_01 (Methyl- <i>l</i> -metabolism)	wt	290	34	146	251.0 Ref.			
rs1667627	het	389	45	292	491.5 (1.2–1.9)			
g-IVS1+3323T>C	var	183	21	152	261.7 (1.3–2.3)			
TEP1_08 (Telomerase)	any var	572	66	444	751.6 (1.2–2.0)			
rs1760897	wt	452	51	261	421.0 Ref.			
g-Ex1–222T>C	het	366	41	280	451.3 (1.0–1.6)			
p-S116P	var	75	8	75	121.8 (1.2–2.6)			
IL8RB_01 (IL-8 Receptor, Angiogenesis, CXCR2)	any var	441	49	355	581.4 (1.1–1.7)			
rs1126579	wt	299	35	250	421.0 Ref.			
Ex3+1235T>C	het	414	48	255	430.7 (0.6–0.9)			
AURKA_15/STK6_15b (Cell cycle kinase, spindle poles in mitosis)	var	150	17	84	140.6 (0.5–0.9)			
rs8173	any var	564	65	339	580.7 (0.6–0.9)			
g-Ex11–347G>C	wt	395	46	224	391.0 Ref.			
FZD7_10 (Wnt receptor, decreases function of antigen presenting cells)	het	356	42	252	441.2 (1.0–1.6)			
rs13034206	var	99	12	92	161.6 (1.2–2.3)			
g-Ex1–1926C>T	any var	455	54	344	611.3 (1.1–1.6)			
MET_01 (Oncogene, hepatocyte growth factor receptor)	wt	453	51	274	441.0 Ref.			
rs41736	het	363	41	272	441.3 (1.0–1.6)			
g-Ex20+60C>T	var	73	8	70	111.7 (1.1–2.4)			
p-D1304D	any var	436	49	342	551.3 (1.1–1.7)			
IGF1_11 (Somatomedin, growth factor)	wt	240	28	198	341.0 Ref.			
rs5742629	het	433	50	291	490.8 (0.6–1.0)			
g-IVS2+12158A>G	var	192	22	102	170.6 (0.5–0.9)			
FZD7_16 (Wnt receptor, decreases function of antigen presenting cells)	any var	625	78	393	660.7 (0.6–0.9)			
rs12474408	wt	492	57	294	501.0 Ref.			
g-*2710A>G	het	325	38	250	421.3 (1.0–1.6)			
MSH6_01 (DNA mismatch recognition)	var	48	6	47	81.7 (1.1–2.6)			
rs3136228	any var	373	44	297	501.4 (1.1–1.7)			
g–356G>T	wt	451	52	269	461.0 Ref.			
GSK3B_45 (Kinase energy metabolism)	het	352	41	261	441.3 (1.0–1.6)			
rs3755557	var	62	7	60	101.7 (1.1–2.5)			
–2675T>A	any var	414	48	321	541.3 (1.1–1.6)			
SLC4A2_04 (Solute exchange)	wt	319	37	192	331.0 Ref.			
rs13240966	het	376	44	248	421.1 (0.9–1.4)			
g-IVS1–194C>G	var	169	20	150	251.6 (1.2–2.1)			
	any var	545	64	398	671.3 (1.0–1.6)			
	wt	539	62	319	541.0 Ref.			
	het	266	31	222	381.5 (1.2–1.9)			
	var	58	7	49	81.3 (0.8–2.0)			
	any var	324	38	271	461.5 (1.2–1.8)			
	wt	514	60	396	671.0 Ref.			
	het	312	36	177	300.8 (0.6–1.0)			
	var	35	4	18	30.6 (0.3–1.1)			
	any var	347	40	195	330.7 (0.6–0.9)			

SNP (function)	Genotype	Controls		Cases		Bladder cancer OR (95% CI)	
		N	%	N	%		%
MBL2_30 (Innate immunity) rs2099902 g.Ex4-710G>A	wt	457	54	346	591.0 Ref.		
	het	329	38	221	370.9 (0.7-1.1)		
	var	66	8	23	40.5 (0.3-0.8)		
	any var	395	46	244	410.8 (0.6-1.0)		
TERT_15 (Telomerase reverse transcriptase) rs13167280 g.IV83-24T>C	wt	659	77	422	711.0 Ref.		
	het	182	21	152	261.4 (1.1-1.8)		
	var	16	2	17	31.7 (0.8-3.4)		
	any var	198	23	169	291.4 (1.1-1.8)		

Analyses were adjusted for gender, age, smoking. Genotype data were missing for ADH1C_14 on 5 controls, 28 cases; MTHFD2_01 3 controls, 27cases; TERT_08 0 controls, 1 case; IL8RB_01 30 controls, 28 cases; AURKA_15 43 controls, 49 cases; FZD7_10 4 controls, 1 case; MET_01 28 controls, 26 cases; IGF1_11 28 controls, 26 cases; FZD7_16 28 controls, 27 cases; MSH6_01 29 controls; 27 cases; GSK3B_45 30 controls, 27 cases; SLC4A2_04 32 controls, 26 cases; MBL2_30 41 controls, 27 cases; TERT_15 36 controls, 26 cases

wt wildtype, het heterozygous, var variant, any variant heterozygous and variant, g. gene, Ex exon, p. protein

Table 3
Individual SNPs selected for association with bladder cancer risk by Kullback–Leibler distance

Function	SNP	Genotype	Controls		Cases		Bladder cancer	
			N	%	N	%	OR (95% CI)	%
Detoxification	GSTZ1_02 g.Ex5-12G>A p.G42R	wt	636		86		801.0 Ref.	
		het	99	14	1	15	191.5 (1.1–2.1)	11.4 (0.4–5.3)
		var	5	1	1	15	201.5 (1.1–2.0)	531.0 Ref.
Aldo/keto reductase PAH metabolism	AKR1C3_35 g.*12259G>A	any var	104		37		411.3 (1.0–1.6)	
		wt	509	59	4	38	61.5 (0.9–2.4)	471.3 (1.0–1.6)
		het	315	41	53	37	511.0 Ref.	
Pigmentation Melanin production	TYR_02 g.IVS3-6895A>G	any var	355		37		431.2 (1.0–1.5)	
		wt	463	37	9	36	60.7 (0.5–1.2)	491.1 (0.9–1.4)
		het	322	46	44	235	401.0 Ref.	
Lipid metabolism	SCARB1_03 g.IVS1-18462G>A	any var	80		12		111.0 (0.7–1.4)	
		wt	402	44	6	545	921.0 Ref.	
		het	383	44	6	46	81.5 (1.0–2.3)	
Metabolism	SLC23A1_05 g.Ex8+22G>A p.V264M	any var	101		–		–	
		wt	482	56	6	46	–	
		het	813	94	6	46	–	
		var	50	–	6	46	–	
		any var	2	6	6	46	–1.5 (0.9–2.2)	

Analyses were adjusted for gender, age, smoking. Genotype data were missing for GSTZ1_02 on 153 controls, 158 cases; AKR1C3_35 on 29 controls, 29 cases; TYR_02 on 28 controls, 27 cases; SCARB1_03 on 28 controls, 26 cases; SLC23A1_05 on 28 controls, 26 cases

wt wildtype, het heterozygous, var variant, any variant heterozygous and variant, g, gene, Ex exon, p, protein

Table 4
Individual SNPs selected for relationship with bladder cancer survival

Function	SNP	Genotype	Cases		Deaths		Bladder cancer HR (95% CI)		Log rank test p value
			N	%	%	N	%		
Surface antigen	CD80_04 rs9282638 g.IVS2-56G>A	wt	321	77	116	681.0 (ref)	0.008		
		het	93	22	54	311.9 (1.4-2.8)			
		var	5	1	1	11.5 (0.2-10.9)			
		any var	98	23	55	321.9 (1.4-2.7)			
Apoptosis	BCL2L1_03 rs1994251 g.IVS2+22130A>C	wt	265	63	85	501.0 (ref)	0.01		
		het	133	32	76	451.5 (1.1-2.1)			
		var	21	5	8	51.5 (0.7-3.3)			
		any var	154	37	84	501.5 (1.1-2.1)			
Detoxify/activate polycyclic aromatic hydrocarbons	EPHX1_15 rs2854461 g.-4786A>C	wt	164	14	85	501.0 (ref)	0.02		
		het	198	47	74	430.8 (0.5-1.1)			
		var	57	14	13	70.5 (0.2-0.9)			
		any var	255	61	87	500.7 (0.5-1.0)			
DNA repair	ERCC4_01 rs1800067 g.Ex8+31G>A p.R415Q	wt	371	89	143	831.0 (ref)	0.1		
		het	45	11	27	161.7 (1.1-2.8)			
		var	3	1	2	15.6 (1.4-23.4)			
		any var	48	12	29	171.8 (1.2-2.9)			

Analyses were adjusted for gender, age, smoking, treatment, tumor stage, grade, tumor size. Genotype data were missing for CD80_04 27 cases; BCL2L1_0329 cases; EPHX1_15 26 cases; ERCC4_01 26 cases

wt wildtype, het heterozygous, var variant, any variant heterozygous and variant, g. gene, Ex exon, p. protein

Table 5
SNPs in genes selected by gene set analysis for relationship with bladder cancer survival

Function	SNP	Genotype	Cases		Deaths		Bladder cancer HR (95% CI)		Log rank test p value
			N	%	N	%	N	%	
Transcription factor Growth control Epithelial cell differentiation	GATA3_29 rs528778 g:IVS4+582C>T	wt	261	62	104	61	611.0 (ref)		<0.001
		het	150	34	51	34	300.9 (0.67–1.3)		
		var	8	2	16	2	92.8 (1.6–5.0)		
		any var	158	36	67	36	391.1 (0.8–1.5)		
Interleukin 8 receptor	IL8RB_01 rs1126579 g:Ex3+1235T>C	wt	194	47	56	47	331.0 (ref)		0.003
		het	173	42	82	42	481.7 (1.2–2.5)		
		var	50	12	34	12	201.7 (1.1–2.8)		
		any var	233	34	116	34	681.7 (1.2–2.5)		
Mitotic kinase Spindle formation	PLK1_15 rs40076 g:IVS3+26A>G	wt	260	62	91	62	531.0 (ref)		0.1
		het	138	33	70	33	411.2 (0.8–1.7)		
		var	21	5	10	5	61.0 (0.5–2.1)		
		any var	159	38	80	38	471.2 (0.8–1.6)		
Glucuronidation Lipophilic molecules	UGT1A1_24 rs1042640 g:Ex5–402G>C	wt	270	65	93	65	551.0 (ref)		0.1
		het	123	30	67	30	401.4 (1.0–2.0)		
		var	22	5	8	5	51.4 (0.6–3.1)		
		any var	155	35	75	35	451.4 (1.0–2.0)		
Progesterone receptor	PGR_05 rs613120 g:IVS2–1671T>C	wt	120	29	34	29	201.0 (ref)		0.005
		het	207	50	81	50	471.2 (0.8–1.9)		
		var	91	22	57	22	331.4 (0.9–2.3)		
		any var	298	77	138	77	801.3 (0.8–2.0)		
Mitotic kinase Chromosome distribution	AURKA/STK6_06 rs6024840 g:IVS7–80T>C	wt	176	56	68	56	471.0 (ref)		0.1
		het	123	39	70	39	481.4 (1.0–2.1)		
		var	15	5	8	5	51.1 (0.5–2.5)		
		any var	138	44	78	44	531.4 (1.0–2.0)		
Telomerase	TERT_02 rs6024840 g:IVS7–80T>C	wt	158	39	82	39	481.0 (ref)		0.1
		het	197	48	75	48	440.8 (0.6–1.1)		
		var	55	13	15	13	80.5 (0.3–1.0)		
		any var	252	61	90	61	520.8 (0.5–1.0)		
Apoptosis	CASP9_01 rs1052576 g:Ex5+32A>G p.Q221R	wt	203	50	55	50	341.0 (ref)		0.003
		het	148	36	77	36	461.3 (0.9–2.0)		
		var	59	14	35	14	211.4 (0.8–2.4)		
		any var	207	50	112	50	671.3 (0.9–2.0)		

Analyses were adjusted for gender, age, smoking, treatment, tumor stage, grade, tumor size. Genotype data were missing for GATA3_29 27 cases; IL8RB_01 28 cases; PLK1_15 27 cases; UGT1A1_24 34 cases; PGR_05 27 cases; AURKA 157 cases; TERT_02 35 cases; CASP9_01 40 cases

wt wildtype, het heterozygous, var variant, any variant heterozygous and variant, g. gene, Ex exon, p. protein