IJC
International Journal of Cancer

Comprehensive analyses of DNA repair pathways, smoking and bladder cancer risk in Los Angeles and Shanghai

Roman Corral[1,2], Juan Pablo Lewinger[1,2], David Van Den Berg[1,2], Amit D. Joshi[3], Jian-Min Yuan[4], Manuela Gago-Dominguez[1,2,5], Victoria K. Cortessis[1,2,6], Malcolm C. Pike[1,2,7], David V. Conti[1,2], Duncan C. Thomas[1,2], Christopher K. Edlund[2], Yu-Tang Gao[8], Yong-Bing Xiang[8], Wei Zhang[8], Yu-Chen Su[1,2] and Mariana C. Stern[1,2]

[1] Department of Preventive Medicine, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA
[2] Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA
[3] Department of Epidemiology, Harvard School of Public Health, Boston, MA
[4] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh Cancer Institute, University of Pittsbugh, Pittsburgh, PA
[5] Galician Foundation of Genomic Medicine, Complexo Hospitalario Universitario Santiago, Servicio Galego de Saude, IDIS, Santiago de Compostela, Spain
[6] Department of Obstetrics and Gynecology, Keck School of Medicine of University of Southern California, Los Angeles, CA
[7] Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY
[8] Department of Epidemiology, Shanghai Cancer Institute and Cancer Institute of Shanghai Jiaotong University, Shanghai, China

Tobacco smoking is a bladder cancer risk factor and a source of carcinogens that induce DNA damage to urothelial cells. Using data and samples from 988 cases and 1,004 controls enrolled in the Los Angeles County Bladder Cancer Study and the Shanghai Bladder Cancer Study, we investigated associations between bladder cancer risk and 632 tagSNPs that comprehensively capture genetic variation in 28 DNA repair genes from four DNA repair pathways: base excision repair (BER), nucleotide excision repair (NER), non-homologous end-joining (NHEJ) and homologous recombination repair (HHR). Odds ratios (ORs) and 95% confidence intervals (CIs) for each tagSNP were corrected for multiple testing for all SNPs within each gene using pACT and for genes within each pathway and across pathways with Bonferroni. Gene and pathway summary estimates were obtained using ARTP. We observed an association between bladder cancer and *POLB* rs7832529 (BER) ($p_{ACT} = 0.003$; $p_{pathway} = 0.021$) among all, and SNPs in *XPC* (NER) and *OGG1* (BER) among Chinese men and women, respectively. The NER pathway showed an overall association with risk among Chinese males (ARTP NER $p = 0.034$). The *XRCC6* SNP rs2284082 (NHEJ), also in LD with *SREBF2*, showed an interaction with smoking (smoking status interaction $p_{gene} = 0.001$, $p_{pathway} = 0.008$, $p_{overall} = 0.034$). Our findings support a role in bladder carcinogenesis for regions that map close to or within BER (*POLB, OGG1*) and NER genes (*XPC*). A SNP that tags both the *XRCC6* and *SREBF2* genes strongly modifies the association between bladder cancer risk and smoking.

Key words: bladder cancer, smoking, DNA repair, *POLB*, *XRCC6*

Abbreviations: BER: Base excision repair; CI: confidence interval; df: degrees of freedom; DNA: deoxyribonucleic acid; HRR: homologous recombination repair; NHW: non-Hispanic White; NOC: N-nitroso compound; NER: nucleotide excision repair; NHEJ: non-homologous end-joining; OR: odds ratio; ROS: reactive oxygen species; SNP: single nucleotide polymorphism; tagSNP: haplotype tagging SNP

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Cancer Institute; **Grant numbers:** 1P01CA86871, 1R01CA065726, 1R01CA114665, P30CA014089; **Grant sponsor:** NIH; **Grant numbers:** T32GM067587, 1R01 ES019876; **Grant sponsor:** NIEHS; **Grant number:** 5P30 ES07048; **Grant sponsor:** FIS Intrasalud; **Grant number:** PS09/02368; **Grant sponsor:** Botin Foundation

DOI: 10.1002/ijc.28693

History: Received 29 Apr 2013; Accepted 11 Nov 2013; Online 2 Jan 2014

Correspondence to: Dr. Mariana C. Stern, USC/Norris Comprehensive Cancer Center, 1441 Eastlake Avenue. Room 5421A. Los Angeles, CA 90089, USA. Tel.: 323-865-0811, E-mail: marianas@usc.edu

Urinary bladder cancer is among the 10 most common cancers worldwide, with its age standardized incidence rate varying by gender and world regions.[1] In Los Angeles County non-Hispanic white men have the highest incidence rate of bladder cancer, followed by Hispanic, African-American and Asian-American men, in spite of comparable profiles of tobacco use. Women show a similar pattern of incidence rates by race, although the overall rates are much lower than men.[2] Chinese from Shanghai have about two-third the incidence rate of bladder cancer of Chinese in Los Angeles.[3] Cigarette smoking and occupational exposure to arylamines are the main established risk factors.[4] Tobacco smoking contributes upwards of 50% of bladder cancer occurrence in men and 20% in women[5]; although more recent data suggests that in the United States the population attributable risk of smoking among men and women might now be comparable.[6] In addition to smoking and occupational exposure to arylamines,[7] use of hair dyes has been identified as a bladder cancer risk factor.[8]

Chemical carcinogens present in tobacco smoke, such as polycyclic aromatic hydrocarbons, aromatic amines, heterocyclic amines and N-nitroso compounds and arylamines from other sources, can induce DNA damage in

What's new?

DNA repair plays a vital role in maintaining DNA integrity in bladder epithelial cells exposed to carcinogens from tobacco smoke. As a result, genetic variations in DNA repair genes could modify bladder cancer risk. Here, analysis of 28 genes that participate in four DNA repair pathways suggests that certain variants in base excision repair and nucleotide excision repair genes may contribute to bladder cancer formation specifically in Chinese populations. Gene-by-environment interaction analyses that included non-Hispanic whites and Chinese suggest that double strand breaks might be the most detrimental type of tobacco-induced DNA damage leading to bladder cancer.

the bladder epithelium.[9] In addition, reactive oxygen species (ROS) present in tobacco smoke,[10] and also generated as a by-product of chemical carcinogen metabolism,[11,12] can contribute to additional DNA damage. Altogether, chemical carcinogens and ROS can contribute to the accumulation of bulky adducts, single (SSB) and double strand breaks (DSB) and various forms of nucleotide base modification or loss which can lead to genomic instability. Modified or lost bases and SSBs are generally repaired through the base excision repair pathway (BER). DSBs are repaired by either the non-homologous end joining (NHEJ) or the homologous recombination repair (HRR) pathways. Bulky adducts are repaired by the nucleotide excision repair (NER) pathway.

Given the important role DNA repair pathways play in maintaining DNA integrity, it has been postulated that inter-individual genetic variation in these pathways may modify bladder cancer risk. Consistent with this hypothesis, individuals with reduced DNA repair proficiency were reported to have higher risk of developing bladder cancer.[13] Several epidemiological studies have investigated the bladder cancer associations with candidate polymorphisms in selected DNA repair genes, and a large pooled and meta-analysis of most of these studies offered support for a role of selected DNA repair variants in bladder carcinogenesis.[14] More recently, a comprehensive analysis of the NER pathway was conducted which offered further support for a role for DNA repair variants in bladder cancer risk.[15]

In this study, we report findings from an extensive pathway-based examination of 632 haplotype-tagging SNPs selected to examine common variation in coding and non-coding regions across 27 DNA repair-related genes, belonging to four DNA repair pathways: BER (*APEX1, LIG3, NEIL1, OGG1, PARP1, POLB* and *XRCC1*), NER (*ERCC1, ERCC2, ERCC4, ERCC5, LIG1, POLD1, XPA* and *XPC*), NHEJ (*DCLRE1C, LIG4, PRKDC, XRCC4, XRCC5* and *XRCC6*) and HRR (*MRE11A, NBN, RAD50, RAD51, RAD52, XRCC2* and *XRCC3*). We conducted these analyses using data from two parallel case-control studies that were similarly designed and carried out in areas of high and low bladder cancer risk: the Los Angeles Bladder Cancer Study and the Shanghai Bladder Cancer Study. We considered the potential modifying role of DNA repair SNPs on the association of gender and smoking with bladder cancer risk.

Material and Methods

Study population

Study participants were enrolled as part of two population-based case-control studies of transitional cell carcinoma of the urinary bladder conducted in Los Angeles County, California, USA and Shanghai, China. Characteristics of the Los Angeles Bladder Cancer (LABC) and Shanghai Bladder Cancer (SBC) studies have been described previously.[16,17] Briefly, the Los Angeles County Cancer Surveillance Program was used to identify cases diagnosed with histologically confirmed bladder cancer, among non-Asian cases between the ages of 25 and 68 years of age from 1987 through 1996. Using a standard procedure,[16] controls were identified among residents of the cases' neighborhoods of residence and individually (1:1) matched to cases by gender, race/ethnicity and age ($\pm$5 years). In Shanghai, the Shanghai Cancer Registry was used to identify cases diagnosed with histologically confirmed bladder cancer, residents of the city of Shanghai and between the ages of 25 and 74 years of age from 1995 to 1998. A previously described algorithm was used to randomly identify population-based controls from within the city of Shanghai,[18] who were frequency matched to bladder cancer cases by gender and five year age groups. In-person questionnaires administered to all study participants were used to collect demographic, lifestyle and medical characteristics up to the reference date, which in Los Angeles was defined for each case-control pair as two years before the case's diagnosis and in Shanghai was defined as two years prior to diagnosis for cases and two years prior to interview for controls. Mean time interval between bladder cancer diagnosis and interview was 11 months for bladder cancer cases in Los Angeles County, and 7 months for bladder cancer cases in Shanghai.[16,17] Blood specimens were collected at the time of interview. Analyses in the current study were restricted to 936 non-Hispanic Whites (NHW) from Los Angeles County (456 cases and 480 controls) and 1,056 Han Chinese from Shanghai (532 cases and 524 controls) with DNA and questionnaire data. The study was approved by Institutional Review Boards at the University of Southern California, the Shanghai Cancer Institute and the University of Pittsburgh.

Tagging SNP selection

Tagging SNPs (tagSNPs) for each DNA repair gene region were selected using Snagger[19] based on the HapMap CEPH

(Utah residents with Northern and Western European Ancestry (CEU)) population and Han Chinese in Beijing, China (CHB), population using data from HapMap release 21, July 2006. TagSNPs were selected using the following criteria: minor allele frequency (MAF) $\geq$ 5%, pairwise $r^2 \geq 0.80$ and a distance from the closest SNP greater than 60 base pairs on the Illumina platform. For each gene, the 5′-UTR- and 3′-UTR regions were extended to include SNPs ~20 kb upstream and ~10 kb downstream. In regions of no or low LD, tagSNPs with a MAF $\geq$5% at a density of ~1 per kb were selected from either HapMap or dbSNP. Finally, nonsynonymous tagSNPs and selected investigator selected SNPs were included regardless of the MAF. With the tagging approach used we were able to capture on average 95.6% (range from 83% to 100%) of genetic variation in CEU and 96.2% (range from 85% to 100%) in CHB, when considering the HapMap release 21, July 2006. This coverage is likely to be lower if we considered the more recent 1,000 Genomes as reference panel.

SNP genotyping and quality control

Peripheral blood lymphocytes were subjected to proteinase K digestion, phenol-chloroform extraction and ethanol precipitation for the purpose of DNA extraction. SNPs were genotyped on the Illumina GoldenGate BeadArray genotyping platform[20] (Illumina, Inc., San Diego, CA, USA) at the Genomics Core of the USC Norris Comprehensive Cancer Center. The Bead Studio software program was used to cluster and call genotypes according to standard Illumina protocols. In addition to Illumina QC measures, cases and controls were mixed on genotyping plates and blinded duplicate samples were included. The observed concordance for duplicate samples was >99%. Genotype data from 30 CEPH trios (Coriell Cell Repository, Camden, NJ) was also used to confirm genotyping reliability and reproducibility. TagSNPs were excluded if more than three discordant genotypes were found in comparison with genotypes from the International HapMap Project.

Further stringent criteria were applied to ensure quality genotyping data. We required that all SNPs have call rates $\geq$0.90 for the combined LABC-SBC study after eliminating SNPs which failed completely. Of the 632 SNPs, five SNPs were eliminated due to call rates of 0%. Supporting Information Table 1 describes all 627 SNPs in this study, including their minor allele frequencies among NHW and Chinese control populations. Analyses that stratified on race were restricted to SNPs with MAF $\geq$5% among Los Angeles controls (545 SNPs) or SNPs with MAF $\geq$5% among Shanghai controls (542 SNPs). Combined analyses of LABC and SBC were restricted to SNPs with MAF $\geq$5% among controls from both study sites (469 SNPs). We required all individuals had overall call rates $\geq$90% and excluded from analyses 192 individuals with overall call rates less than 90%. After excluding subjects with call rates less than 90%, we had genotyping results available for 1,800 individuals out of a total of 1,992. Individuals with genotyping data did not differ significantly from those without genotyping data for key characteristics, such as those listed in Table 1.

Deviations of observed genotype frequencies from those expected under Hardy-Weinberg equilibrium (HWE) were examined among Los Angeles and Shanghai controls separately using exact tests. The p-value when testing deviations of observed genotype frequencies from those expected under HWE was deemed significant if $p < 0.00008$ using exact tests (Bonferroni-corrected p-value; $\alpha = 0.05/627$). We did not observe evidence of deviations of observed from expected values among Los Angeles non-Hispanic white controls or Shanghai Chinese controls.

Statistical analysis

SNP main effects. In order to include all available individuals in our study, regardless of availability of 1:1 matched controls, we grouped individuals according to their reference age (<45, 45–49, 50–54, 55–59 and $\geq$60 years for Los Angeles non-Hispanic whites and <45, 45–49, 50–54, 55–59, 60–64 and $\geq$65 years for Shanghai Chinese), gender and study site and used it to group individuals in conditional logistic regression models used to estimate relative risks with odds ratios (ORs) and 95% confidence intervals (95% CI). Assuming a log-additive mode of action, we estimated per-allele ORs and 95% CI for the associations between each tagSNP and bladder cancer. Models were adjusted for smoking status (never/quit/current) in the reference year. Analyses were conducted separately by study site and jointly with adjustment for study site; we assessed for potential heterogeneity of SNP main effects across both study sites using likelihood ratio tests. Given the observed disparities in bladder cancer incidence between males and females, both in Los Angeles and Shanghai, we hypothesized that different environmental risk factors could associate with each gender. If some of these risk factors contribute to bladder carcinogenesis through the accumulation of DNA damage, we speculated that we could observe different associations between DNA repair SNPs and bladder cancer for males and females. To test this hypothesis, we assessed potential heterogeneity of SNP main effects by gender using likelihood ratios tests.

Multiple testing was conducted in a hierarchical bottom-up manner. We first corrected for multiple SNP tests within each gene region, then for multiple genes within the corresponding DNA repair pathway, and finally across all four DNA repair pathways investigated. Specifically, for each SNP within each gene region, crude p-values (p_{crude}) were corrected for multiple testing using the P_{ACT} (p-value adjusted for correlated tests) approach implemented within R.[21] We corrected for overall significance across gene regions within each pathway ($p_{pathway}$) using a Bonferroni correction of the P_{ACT} corrected p-value. Finally, we further corrected for overall statistical significance across all four investigated pathways ($p_{overall}$) using a Bonferroni correction of the pathway specific ($p_{pathway}$) p-value.

Table 1. Characteristics of non-Hispanic whites in Los Angeles County and Shanghai Chinese

	Los Angeles County			Shanghai		
	Cases $n = 456$	Controls $n = 480$		Cases $n = 532$	Controls $n = 524$	
Mean age at enrollment (SD)	56 (±7)	56 (±8)		63 (±10)	64 (±10)	
Age at enrollment (y)						
≤45	51 (11%)	62 (13%)		51 (10%)	43 (8%)	
45–49	54 (12%)	57 (12%)		30 (6%)	17 (3%)	
50–54	83 (18%)	93 (19%)		38 (7%)	25 (5%)	
55–59	138 (30%)	128 (27%)		43 (8%)	63 (12%)	
60–64	129 (28%)	104 (22%)		137 (26%)	116 (22%)	
>65	1 (0%)	35 (7%)	<0.001	233 (44%)	260 (50%)	0.016
Gender						
Male	357 (78%)	374 (78%)		421 (79%)	404 (77%)	
Female	99 (22%)	106 (22%)	0.890	111 (21%)	120 (23%)	0.424
Smoking status						
Never	83 (18%)	183 (38%)		178 (33%)	233 (44%)	
Former	173 (38%)	212 (44%)		75 (14%)	84 (16%)	
Current	200 (44%)	85 (18%)	<0.001	279 (52%)	207 (40%)	<0.001
Smoking intensity (cigarettes/day)						
Never	83 (18%)	183 (38%)		178 (33%)	233 (44%)	
<20	75 (16%)	88 (18%)		164 (31%)	155 (30%)	
≥20	298 (65%)	209 (44%)	<0.001	190 (36%)	136 (26%)	<0.001
Smoking duration (years)						
Never	83 (18%)	183 (38%)		178 (33%)	233 (45%)	
<29	162 (36%)	190 (40%)		106 (20%)	101 (19%)	
≥29	211 (46%)	107 (22%)	<0.001	248 (47%)	190 (36%)	0.001
Pack-years of smoking						
Never	83 (18%)	183 (38%)		178 (33%)	233 (44%)	
<24 pack-years	116 (25%)	151 (31%)		155 (29%)	142 (27%)	
≥24 pack-years	257 (56%)	146 (30%)	<0.001	199 (37%)	149 (28%)	0.001

Pathway analyses. In order to capture gene and pathway level effects that may not be detectable through any single SNP, we performed gene-based and pathway-based tests using the Adaptive Rank-Truncated Product (ARTP) method.[22] ARTP adaptively combines single SNP *p*-values within a gene region or a pathway to obtain a single test statistic for the gene or pathway and assesses significance of the test *via* a permutation procedure. Unlike a multiple testing procedure like P_{ACT}, which accounts for multiple SNP tests in order to properly control the type I error, ARTP combines information across SNPs within a gene or a pathway in order to increase the power to detect a gene or pathway level effect.

SNP-smoking interactions. We investigated SNP-smoking interactions considering the following smoking variables: smoking status (never, former, current), smoking intensity (never, <20, ≥20 cigarettes per day), smoking duration (never, <29, ≥29 years of smoking) and pack-years of smoking (never, <24, ≥24 pack-years). Three-level variables were generated using the median value among smoking controls as a cut point for cigarettes per day, years of smoking and pack-years. Interactions between SNPs and exposures were investigated on a multiplicative scale using conditional logistic models, assuming a log-additive mode of risk and using likelihood ratio tests that included product terms between each tagSNP and a three-level exposure variable coded with dummy variables. Tests of trend across categories of exposure were conducted by assigning median values to every tertile of exposure and modeling the categories as continuous. Interaction between SNPs and smoking status assumed that smoking status (never = 0, quit = 1, current = 2) was a categorical variable in the interaction model, while the *p*-values for trend

were calculated assuming smoking status as continuous in the interaction model.

Similar to our hierarchical approach for multiple testing correction for SNP main effects, within each gene region, crude interaction *p*-values for each SNP (interaction p_{crude}) were adjusted using a Bonferroni correction (P_{ACT} supports multiple tests of SNP main effects but not multiple tests of SNP by exposure interactions) that considered the number of SNPs investigated within each corresponding gene region (interaction p_{gene}). These corrected interaction p-values were further adjusted using a Bonferroni correction for the number of gene regions within each specific pathway (interaction $p_{pathway}$). Finally, these corrected interaction *p*-values were further adjusted using Bonferroni for pathway-wide significance (interaction $p_{overall}$), considering that a total of four pathways had been investigated. In all levels of correction, statistical significance was declared if corrected *p*-values were <0.05. All statistical tests conducted were two sided and all analyses were performed using Stata 11/SE (Stata Corporation, College Station, TX) and the statistical package R 2.15 (The R Project for Statistical Computing, http://www.r-project.org).

Results

Characteristics of cases and controls are summarized in Table 1. Briefly, males accounted for approximately 80% of study participants in both Los Angeles County and Shanghai. Mean age at enrollment for cases was 56 years of age in Los Angeles County and 64 years of age in Shanghai. While 44% of Shanghai cases were older than 65 years of age, less than 1% of Los Angeles cases were older than 65 years of age. Reported rates of cigarette smoking were higher among Los Angeles County cases and controls than among Shanghai cases and controls.

DNA repair SNPs and bladder cancer risk

We investigated associations between DNA repair tagSNPs and bladder cancer risk among NHW from the LABCS and Chinese from the SBCS, separately and combined. Among the 545 tagSNPs investigated among NHW in the LABC study 21 showed statistically significant associations with bladder cancer ($p_{crude} < 0.05$); however, none remained significant after within gene region correction ($p_{ACT} > 0.05$). None of these 21 tagSNPs showed statistically significant associations among Shanghai Chinese (Supporting Information Table 1).

Among the 542 tagSNPs investigated among Shanghai Chinese, 30 tagSNPs were statistically significantly associated with bladder cancer ($p_{crude} < 0.05$), and five of them remained statistically significant after multiple comparisons adjustment within gene region ($P_{ACT} < 0.05$): one in the *POLB* gene (rs7832529, OR = 1.5; 95% CI = 1.2–1.9; $p_{ACT} = 0.003$), one in the *POLD1* gene (rs2244095, OR = 0.8; 95%CI = 0.6–0.9; $p_{ACT} = 0.025$) and three in the *XPC* gene (rs2607734, OR = 1.3, 95% CI = 1.1–1.6, $p_{ACT} = 0.020$; rs2279017, OR = 1.3; 95%CI = 1.1–1.6, $p_{ACT} = 0.024$; rs2228001, OR = 1.3, 95%CI = 1.1–1.6, $p_{ACT} = 0.028$) (Table 2).

Among the 469 tagSNPs investigated among the LABC and SBC combined, 24 tagSNPs showed statistically significant associations with bladder cancer ($p_{crude} < 0.05$). Only 3 tagSNPs—the same ones we observed to be associated among Shanghai Chinese from the SBCS—remained statistically significant after multiple comparisons adjustment within gene region ($p_{ACT} < 0.05$): one in the *POLB* gene (rs7832529, OR = 1.5, 95% CI = 1.2–1.9, $p_{ACT} = 0.003$) and two in the *POLD1*gene (rs2244095, OR = 0.8, 95% CI = 0.7–0.9, $p_{ACT} = 0.018$; rs2546551, OR = 0.8, 95% CI = 0.7–0.9, $p_{ACT} = 0.049$) genes (Table 2). Of these three SNPs, only one remained statistically significant when correcting for all genes within the corresponding pathway (BER) and showed a borderline significant association when correcting for all pathways considered (*POLB* rs7832529 $p_{ACT} = 0.003$; $p_{pathway} = 0.021$; $p_{overall} = 0.084$). None of these three tagSNPs showed statistically significant heterogeneity by racial groups (NHW versus Chinese); results among Chinese and NHW were of similar magnitude and direction but were statistically significant only among Chinese. Conversely, the three tagSNPs in the *XPC* gene found to be statistically significantly associated with bladder cancer risk among Chinese showed heterogeneity by race (rs2607734 heterogeneity $p = 0.041$; rs2279017 heterogeneity $p = 0.044$; rs2228001 heterogeneity $p = 0.058$), with the association being restricted to Chinese.

DNA repair SNPs and smoking interactions

We conducted gene by smoking interaction analyses among NHW and Chinese combined. None of the SNPs previously identified to associate with bladder cancer risk (Table 2) were found to modify the risk of smoking on bladder cancer. *XRCC6* (rs2284082), *XPA* (rs7853179), *XRCC3* (rs709400) and *DCLRE1C* (rs1079622) were found to modify the effect of smoking across different measures of exposure, with interaction test *p*-values that achieved statistical significance within each gene, but not at the pathway level (Table 3). The only exception was *XRCC6* SNP rs2284082 (NHEJ pathway), which showed an interaction that achieved within gene region and within pathway and overall pathway statistical significance (Table 4). Specifically, among carriers of one (CT) or two (CC) copies of the major allele C, statistically significant trends were observed for the associations between smoking pack-years, years of smoking, cigarettes per day and smoking status, with greater strengths of association for CC carriers than CT carriers. Instead, among carriers of two copies of the minor allele T (TT), non-statistically significant positive trends, with reduced estimates, were observed (Table 4). For all smoking variables considered, except cigarettes per day, tests of interaction remained statistically significant after correction for multiple testing at the gene and pathway levels (Smoking pack-years interaction $p_{gene} = 0.003$, $p_{pathway} = 0.020$; years of smoking interaction $p_{gene} = 0.008$, $p_{pathway} =$

Table 2. SNPs associated with bladder cancer risk in the Los Angeles-Shanghai study

Pathway	Gene	tagSNP	MAF	Ca	Co	OR[1]	LCI	UCI	p_{crude}	p_{ACT}	$p_{pathway}$	$p_{overall}$	p_{Het}
NHW (LABC)													
BER	*POLB*	rs7832529	0.05	351	405	1.4	0.9	2.1	0.186	0.662	1.000	1.000	
NER	*POLD1*	rs2244095	0.11	353	407	0.9	0.6	1.2	0.473	0.884	1.000	1.000	
NER	*POLD1*	rs2546551	0.44	353	404	0.8	0.7	1.0	0.066	0.354	1.000	1.000	
NER	*XPC*	rs2607734	0.43	355	409	1.0	0.8	1.3	0.857	1.000	1.000	1.000	
NER	*XPC*	rs2279017	0.43	354	408	1.0	0.8	1.3	0.873	1.000	1.000	1.000	
NER	*XPC*	rs2228001	0.43	352	409	1.0	0.8	1.3	0.773	1.000	1.000	1.000	
Chinese (SBC)													
BER	*POLB*	rs7832529	0.12	509	518	1.5	1.2	2.0	0.001	0.009	0.060	0.239	
NER	*POLD1*	rs2244095	0.35	513	514	0.8	0.6	0.9	0.004	0.025	0.173	0.693	
NER	*POLD1*	rs2546551	0.16	512	512	0.8	0.6	1.0	0.049	0.219	1.000	1.000	
NER	*XPC*	rs2607734	0.36	514	520	1.3	1.1	1.6	0.002	0.020	0.141	0.562	
NER	*XPC*	rs2279017	0.36	510	520	1.3	1.1	1.6	0.003	0.024	0.168	0.670	
NER	*XPC*	rs2228001	0.36	513	521	1.3	1.1	1.6	0.004	0.028	0.197	0.788	
NHW & Chinese (LABC & SBC)													
BER	*POLB*	rs7832529		860	923	1.5	1.2	1.9	<0.001	0.003	0.021	0.084	0.564
NER	*POLD1*	rs2244095		866	921	0.8	0.7	0.9	0.003	0.018	0.125	0.500	0.350
NER	*POLD1*	rs2546551		865	916	0.8	0.7	0.9	0.009	0.049	0.342	1.000	0.676
NER	*XPC*	rs2607734		869	929	1.2	1.0	1.4	0.015	0.095	0.667	1.000	0.041
NER	*XPC*	rs2279017		864	928	1.2	1.0	1.4	0.018	0.111	0.778	1.000	0.044
NER	*XPC*	rs2228001		865	930	1.2	1.0	1.4	0.016	0.101	0.709	1.000	0.058

[1]Per allele ORs and 95% CIs estimated from conditional logistic regression models assuming a log-additive mode of risk and adjusting for smoking status in reference year

LCI = 95% lower confidence interval; UCI = 95% upper confidence interval; p_{crude} = unadjusted for multiple testing p-value; p_{ACT} = *p*-value corrected for multiple testing within gene region; $p_{pathway}$ = *p*-value corrected for multiple testing within gene region and within pathway; $p_{overall}$ = *p*-value corrected for testing across all SNPs and pathways; p-Het = LRT *p*-value from test of heterogeneity.

Table 3. DNA repair SNPs × smoking interactions among NHW from Los Angeles County & Shanghai Chinese

Exposure	# SNPs with interaction $p_{crude} < 0.05$	Pathway	Gene	SNP	Interaction p_{crude}	Interaction p_{gene}	Interaction $p_{pathway}$	Interaction $p_{overall}$
Years of smoking	29	NHEJ	*XRCC6*	rs2284082	0.001	0.008	0.046	0.185
		NER	*XPA*	rs7853179	0.002	0.023	0.164	0.656
		HR	*XRCC3*	rs709400	0.003	0.036	0.250	0.999
Pack-years of smoking	26	NHEJ	*XRCC6*	rs2284082	<0.001	0.003	0.020	0.079
		NHEJ	*DCLRE1C*	rs10796227	0.002	0.033	0.199	0.794
Cigarettes per day	8	NHEJ	*XRCC6*	rs2284082	0.015	0.093	0.556	1.000
Smoking Status	35	NHEJ	*XRCC6*	rs2284082	<0.001	0.001	0.008	0.034
		HR	*XRCC3*	rs709400	0.002	0.025	0.177	0.706
		NER	*XPA*	rs7853179	0.003	0.048	0.338	1.000
		NHEJ	*DCLRE1C*	rs10796227	0.004	0.050	0.302	1.000

Table 4. XRCC6 rs2284082 x smoking interactions and bladder cancer risk among NHW from Los Angeles County & Shanghai Chinese

	Cases/Controls			CC				CT				TT				Interaction *p*-values	
Smoking variables	**CC**	**CT**	**TT**	**OR**	**LCI**	**UCI**	**p-value**	**OR**	**LCI**	**UCI**	**p-value**	**OR**	**LCI**	**UCI**	**p-value**		
Smoking pack-yrs																	
Never	67/144	111/174	54/57	1.0				1.0				1.0				p_{crude}	0.001
<24 pack-years	84/112	105/99	45/45	1.9	1.3	2.8	0.001	1.6	1.2	2.2	<0.001	1.4	0.9	2.3	0.151	p_{gene}	0.003
≥24 pack-years	154/91	173/129	45/42	4.3	3.0	6.2	<0.001	2.4	1.9	3.2	<0.001	1.4	0.9	2.2	0.193	$p_{pathway}$	0.020
p for trend							<0.001				<0.001				0.288	$p_{overall}$	0.079
Years of smoking																	
Never	67/144	111/174	54/57	1.0				1.0				1.0				p_{crude}	0.001
<29	90/109	95/104	38/36	1.8	1.2	2.7	0.003	1.5	1.1	2.0	0.007	1.2	0.7	2.1	0.417	p_{gene}	0.008
≥29	148/94	183/124	52/51	4.4	3.0	6.4	<0.001	2.6	2.0	3.4	<0.001	1.5	1.0	2.4	0.077	$p_{pathway}$	0.046
p for trend							<0.001				<0.001				0.077	$p_{overall}$	0.185
Cigarettes per day																	
Never	67/144	111/174	1.0				1.0					1.0				p_{crude}	0.015
<20	81/84	95/94	38/41	2.3	1.5	3.4	<0.001	1.7	1.3	2.2	<0.001	1.2	0.8	2.0	0.405	p_{gene}	0.093
≥20	157/119	183/134	52/46	3.5	2.4	5.0	<0.001	2.3	1.8	3.0	<0.001	1.5	1.0	2.4	0.069	$p_{pathway}$	0.556
p for trend							<0.001				<0.001				0.068	$p_{overall}$	1.000
Smoking Status																	
Never	67/144	111/174	54/57	1.0				1.0				1.0				p_{crude}	<0.001
Former	78/118	87/100	30/31	1.6	1.1	2.4	0.022	1.4	1.1	1.9	0.021	1.3	0.7	2.1	0.394	p_{gene}	0.001
Current	160/85	191/128	60/56	4.7	3.3	6.9	<0.001	2.6	2.0	3.4	<0.001	1.4	0.9	2.3	0.123	$p_{pathway}$	0.008
p for trend							<0.001				<0.001				0.125	$p_{overall}$	0.032

LCI = 95% lower confidence interval; UCI = 95% upper confidence interval; p_{crude} = unadjusted for multiple testing *p*-value; p_{gene} = *p*-value corrected for multiple testing within gene region; $p_{pathway}$ = *p*-value corrected for multiple testing within gene region and within pathway; $p_{overall}$ = *p*-value corrected for testing across all SNPs and pathways.

0.046; smoking status interaction $p_{gene} = 0.001$, $p_{pathway} = 0.008$) (Table 4). Test of interaction for smoking status also remained statistically significant when further correcting for the total number of DNA repair pathways investigated (smoking pack-years interaction $p_{overall} = 0.032$) (Table 4).

DNA repair SNPs by gender interactions

To explore possible heterogeneity of the SNP-bladder cancer associations, we conducted stratified analysis by gender among NHW, Chinese and among both sites combined (Table 5). Among NHW males but not NHW females, we observed inverse associations for three linked *LIG1* SNPs (rs2007183, rs20579 and rs3730912) with bladder cancer that were statistically significant after within-gene-region correction ($p_{ACT} < 0.05$) and showed evidence of heterogeneity by gender ($p_{heterogeneity} < 0.05$) (Table 5).

Among Chinese, we observed three tagSNPs in the *OGG1* gene that showed evidence of statistically significant heterogeneity by gender. These three SNPs were inversely associated with bladder cancer risk only among females, and the associations remained statistically significant after within-gene corrections, and for one of them remained significant after pathway correction as well (rs6809452, OR = 0.5; 95% CI = 0.3–0.8, $p_{ACT} = 0.007$, $p_{pathway} = 0.046$; rs1052133, OR = 0.6, 95% CI = 0.4–0.8, $p_{ACT} = 0.026$; rs2072668, OR = 0.6, 95% CI = 0.4–0.9, $p_{ACT} = 0.049$). Similar estimates were observed among NHW females and among NHW and Chinese females combined, but estimates did not reach statistical significance (data not shown). We also observed that the previously observed associations of the *POLB* tagSNP (rs7832529) and the 3 *XPC* tagSNPs (rs26077734, rs2228001 and rs2279017) with bladder cancer risk among all Chinese individuals combined, plus an additional new *XPC* tagSNP (rs2305843), seemed restricted to males, but tests of heterogeneity were not statistically significant (Table 5).

Similarly, among males in the combined study (NHW and Chinese), three *XPC* tagSNPs (rs2305843 rs2607734 and rs2228001) were statistically significantly associated with bladder cancer risk. In addition, the previously observed association between the *POLD1* tagSNPs (rs2546651 and rs2244095) and bladder cancer risk among both races combined seemed restricted to males. However, for neither of these tagSNPs were tests of heterogeneity by gender statistically significant (Table 5).

Pathway analyses

We used the ARTP approach to obtain a summary *p*-value for the association of each gene and pathway considered in the study with bladder cancer risk (Table 6). Among NHW, only *LIG1* (NER pathway) achieved gene-wide statistical significance among males. Instead, among Chinese, six genes appeared associated with susceptibility to bladder cancer achieving ARTP gene-wide significance, with four of them showing heterogeneity by gender: *OGG1* (Chinese females $p_{ARTP\ gene} = 0.015$), *POLB* (All Chinese $p_{ARTP\ gene} = 0.010$, Chinese males $p_{ARTP\ gene} = 0.048$), *RAD50* (All Chinese $p_{ARTP\ gene} = 0.034$, Chinese males $p_{ARTP\ gene} = 0.023$), *POLD1* (All Chinese $p_{ARTP\ gene} = 0.021$), *XPC* (All Chinese $p_{ARTP\ gene} = 0.017$, Chinese males $p_{ARTP\ gene} = 0.003$) and finally *XRCC6* (All Chinese $p_{ARTP\ gene} = 0.010$, Chinese females $p_{ARTP\ gene} = 0.043$). Three of these genes showed ARTP gene-wide significance when all NHW and Chinese combined: *POLB* (Chinese & NHW $p_{ARTP\ gene} = 0.013$), *RAD50* (Chinese & NHW $p_{ARTP\ gene} = 0.048$), *POLD1* (Chinese & NHW $p_{ARTP\ gene} = 0.013$) and *XPC* (Chinese & NHW $p_{ARTP\ gene} = 0.045$) (Table 6). When considering overall pathway associations, we only observed an association of pathway-wide significance for the NER pathway among Chinese males ($p_{ARTP\ pathway} = 0.034$), and we observed a pathway-wide ARTP *p*-value of borderline significance when considering all Chinese combined ($p_{ARTP\ pathway} = 0.068$) (Table 6).

Discussion

In this study we investigated the association between a comprehensive SNP panel that captured genetic variation in genes that play key roles in four different DNA repair pathways and bladder cancer risk. Our most consistent and key findings were an association between *POLB* rs7832529 and bladder cancer risk, predominantly among Chinese, an association between *OGG1* rs6809452 and bladder cancer risk among Chinese women only and an association between *XPC* rs2607734 and bladder cancer risk among Chinese men only. POLB and OGG1 play key roles in the BER pathway and XPC participates in the NER pathway. Analyses that summarized the effects of all SNPs within each gene, obtained using the ARTP approach for both genders combined confirmed a role for *POLB* in bladder cancer risk among Chinese and also indicated associations between *RAD50* (HRR pathway), *POLD1* (NER pathway), *XPC* (NER pathway), *LIG1* (NER pathway), *OGG1* (BER pathway) and *XRCC6* (NHEJ pathway). However, when considering estimates that summarized the effect of all genes within each of the four pathways, we observed only a statistically significant association for the NER pathway among Chinese males and a borderline statistically significant one among all Chinese combined. When considering cigarette smoking variables we found consistent evidence that the *XRCC6* rs2284082 SNP (NHEJ pathway) modified the effect of smoking. Estimates of interaction for this SNP remained statistically significant after correction for multiple testing within each gene, within the NHEJ pathway and across all four pathways. None of the genes in the other three pathways showed strong evidence of effect modification by smoking. Altogether, these findings suggest that among Chinese, particularly men, there are bladder cancer risk factors, other than smoking, that elicit the BER and NER pathways and may play key roles in bladder cancer formation. Alternatively, they suggest that presence of these genetic variants, may predispose individuals to developing bladder cancer, independently of environmental exposures, perhaps due to loss over time of DNA repair proficiency and inability to repair DNA damage that may accumulate with age.

Table 5. SNPs associated with bladder cancer risk, among males and females, in the Los Angeles-Shanghai study

			Males									Females								
Pathway	Gene	SNP	CA	CO	OR[1]	LCI	UCI	p_{crude}	p_{ACT}	$p_{pathway}$	$p_{overall}$	CA	CO	OR[1]	LCI	UCI	p_{crude}	p_{ACT}	$p_{pathway}$	$p_{overall}$	LRp
NHW																					
NER	*LIG1*	rs2007183	277	322	0.6	0.4	0.9	0.005	0.041	0.288	1.000	76	86	1.7	0.8	3.8	0.165	0.609	1.000	1.000	0.013
NER	*LIG1*	rs20579	279	323	0.6	0.4	0.9	0.006	0.044	0.308	1.000	76	86	1.9	0.9	4.0	0.113	0.487	1.000	1.000	0.008
NER	*LIG1*	rs3730912	279	323	0.6	0.4	0.9	0.008	0.055	0.387	1.000	76	86	1.7	0.7	4.0	0.230	0.656	1.000	1.000	0.027
Chinese																					
NER	*XPC*	rs2607734	407	400	1.5	1.2	1.8	0.001	0.005	0.032	0.127	107	120	1.0	0.7	1.5	0.987	1.000	1.000	1.000	0.092
NER	*XPC*	rs2228001	406	402	1.4	1.2	1.8	0.001	0.008	0.055	0.219	107	119	1.0	0.7	1.5	0.991	1.000	1.000	1.000	0.107
NER	*XPC*	rs2279017	403	401	1.4	1.2	1.8	0.001	0.008	0.056	0.223	107	119	1.0	0.7	1.5	0.902	1.000	1.000	1.000	0.139
NER	*XPC*	rs2305843	406	401	1.4	1.1	1.7	0.004	0.033	0.227	0.909	107	120	1.0	0.7	1.4	0.764	1.000	1.000	1.000	0.098
BER	*POLB*	rs7832529	403	398	1.5	1.1	2.0	0.009	0.057	0.400	1.000	106	120	1.8	1.0	3.2	0.043	0.237	1.000	1.000	0.524
BER	*OGG1*	rs6809452	407	402	0.9	0.7	1.1	0.324	0.720	1.000	1.000	107	120	0.5	0.3	0.8	0.001	0.007	0.046	0.184	0.011
BER	*OGG1*	rs1052133	404	402	0.9	0.8	1.1	0.443	0.773	1.000	1.000	107	119	0.6	0.4	0.8	0.004	0.026	0.179	0.717	0.025
BER	*OGG1*	rs2072668	405	402	0.9	0.8	1.2	0.518	1.000	1.000	1.000	107	118	0.6	0.4	0.9	0.008	0.049	0.342	1.000	0.038
NHW & Chinese																					
BER	*POLB*	rs7832529	678	719	1.4	1.1	1.8	0.015	0.076	0.534	1.000	182	204	2.1	1.3	3.4	0.004	0.024	0.170	0.679	0.145
NER	*XPC*	rs2305843	683	715	1.3	1.1	1.6	0.004	0.030	0.211	0.845	183	204	1.0	0.7	1.4	0.948	1.000	1.000	1.000	0.131
NER	*XPC*	rs2607734	686	723	1.3	1.1	1.5	0.006	0.040	0.282	1.000	183	206	1.0	0.7	1.3	0.984	1.000	1.000	1.000	0.190
NER	*XPC*	rs2228001	683	725	1.3	1.1	1.5	0.006	0.041	0.290	1.000	182	205	1.0	0.7	1.3	0.981	1.000	1.000	1.000	0.178
NER	*POLD1*	rs2546551	683	714	0.8	0.6	0.9	0.007	0.045	0.314	1.000	182	202	0.9	0.6	1.3	0.601	1.000	1.000	1.000	0.419
NER	*POLD1*	rs2244095	683	718	0.8	0.6	0.9	0.009	0.050	0.347	1.000	183	203	0.8	0.5	1.1	0.135	0.390	1.000	1.000	0.932

[1]Per allele ORs and 95% CIs estimated from conditional logistic regression models assuming a log-additive mode of risk and adjusting for smoking status in reference year
LCI = 95% lower confidence interval; UCI = 95% upper confidence interval; p_{crude} = unadjusted for multiple testing p-value; p_{ACT} = p-value corrected for multiple testing within gene region; $p_{pathway}$ = *p*-value corrected for multiple testing within gene region and within pathway; $p_{overall}$ = *p*-value corrected for testing across all SNPs and pathways; LRp = LRT p-value from test of heterogeneity

Table 6. Gene- and pathway-level summary p-values from ARTP pathway analyses in LABCS and SBCS

		ARTP Gene *p*-value							ARTP Pathway *p*-value						
		NHW			Chinese				NHW			Chinese			
Pathway	Gene/ Region	All	Females	Males	All	Females	Males	Combined All	All	Females	Males	All	Females	Males	Combined All
BER									0.746	0.443	0.4845	0.356	0.143	0.329	0.724
	APEX1	0.926	0.932	0.898	0.831	0.920	0.791	0.926							
	LIG3	0.431	0.061	0.904	0.834	0.891	0.767	0.474							
	NEIL1	0.633	0.727	0.712	0.881	0.922	0.724	0.894							
	OGG1	0.674	0.678	0.812	0.225	**0.015**	0.556	0.486							
	PARP1	0.170	0.650	0.066	0.895	0.085	0.761	0.628							
	POLB	0.416	0.248	0.980	**0.010**	0.262	**0.048**	**0.013**							
	XRCC1	0.736	0.252	0.828	0.209	0.887	0.116	0.694							
HRR									0.554	0.587	0.784	0.510	0.772	0.219	0.676
	MRE11A	0.272	0.239	0.285	0.805	0.589	0.849	0.566							
	NBN	0.380	0.310	0.366	0.471	0.718	0.651	0.795							
	RAD50	0.591	0.189	0.871	**0.034**	0.824	**0.023**	**0.048**							
	RAD51	0.214	0.923	0.196	0.930	0.965	0.970	0.402							
	RAD52	0.967	0.488	0.959	0.821	0.622	0.946	0.935							
	XRCC2	0.154	0.267	0.466	0.779	0.155	0.978	0.276							
	XRCC3	0.384	0.872	0.877	0.321	0.395	0.255	0.664							
NER									0.649	0.672	0.234	0.068	0.293	**0.034**	0.107
	ERCC1-ERCC2		0.682	0.824	0.740	0.133	0.161	0.193	0.087						
	ERCC4	0.312	0.2105	0.918	0.069	0.160	0.267	0.487							
	ERCC5	0.177	0.7625	0.181	0.332	0.916	0.462	0.360							
	LIG1	0.139	0.3745	**0.025**	0.716	0.302	0.795	0.942							
	POLD1	0.430	0.4705	0.315	**0.021**	0.057	0.105	**0.013**							
	XPA	0.956	0.1245	0.985	0.915	0.491	0.806	0.886							
	XPC	0.563	0.9705	0.440	**0.017**	0.941	**0.003**	**0.045**							
NHEJ									0.598	0.182	0.698	0.236	0.206	0.774	0.397
	DCLRE1C	0.975	0.190	0.7845	0.728	0.216	0.942	0.717							
	LIG4	0.715	0.241	0.841	0.908	0.711	0.935	0.841							
	PRKDC	0.587	0.352	0.7015	0.789	0.298	0.840	0.540							
	XRCC4	0.104	0.125	0.140	0.174	0.247	0.315	0.129							
	XRCC5	0.768	0.629	0.678	0.292	0.333	0.280	0.595							
	XRCC6	0.821	0.076	0.808	**0.038**	**0.043**	0.284	0.458							

Finally, our findings support a role for the NHEJ pathway in smoking-induced bladder cancer risk, suggesting that among all types of damage induced by tobacco carcinogens, double strand breaks seem to be the ones more detrimental for cancer risk. In support of this, two other NHEJ genes (*DCLRE1C* and *XRCC3*) were also found to modify the effect of smoking, although findings were not as significant as for *XRCC6*.

The number of variants and genes investigated in DNA repair pathways in association with bladder cancer risk has been limited. In collaboration with the International Consortium of Bladder Cancer Studies, we previously published a meta-analysis and pooled analyses of 10 common variants in seven genes and reported that three SNPs (*ERCC2* rs1799793, *NBN* rs1805794 and *XPC* rs2228000) were associated with a modest increase in bladder cancer risk.[14] GWAS, meta-analysis of GWAS and pathway-based analysis of GWAS have identified multiple loci associated with bladder cancer susceptibility in subjects of European ancestry.[23–29] Whereas several SNPs located in carcinogen metabolism enzyme coding genes have achieved genome-wide significance, no SNPs located in DNA repair genes have achieved genome-wide significance to date. We summarize below what is known about

the genetic regions for which we found stronger evidence of an association with bladder cancer risk (*XPC, POLB, OGG1* and *POLD)* and evidence of interaction with smoking (*XRCC6)*.

Our pathway-based analyses point to the NER pathway as relevant for bladder cancer risk. Associations between SNPs in the *XPC* and *POLD1* genes among Chinese seemed to be responsible for the overall observed association with this pathway. NER is involved in the repair of bulky DNA adducts, such as those induced by tobacco smoke carcinogens.[30] The xeroderma pigmentosum complementation group C gene (*XPC*) (HGNC 12816) is located on chromosome 3p25. XPC detects and binds to DNA adducts and initiates recruitment of other NER pathway proteins at the site of damage.[31,32] Our individual SNP analyses and overall gene analyses suggested an association between bladder cancer risk and *XPC*. Pooled analyses of most available epidemiological studies with data on selected *XPC* polymorphisms, including ours, showed an association for *XPC* rs2228000 with bladder cancer risk among NHW and no association with SNP rs2228001.[14] In this study, we could not replicate the association with rs2228000 among NHW or Chinese; however, we report a statistically significant association between *XPC* rs2228001 and bladder cancer risk among Chinese males.[14] The functional relevance/biological mechanism of the variant is unknown. There are two 3′UTR SNPs nearby that have been reported to affect XPC protein expression: rs2470352 and rs2470458[33]; however, neither of these SNPs are in LD with rs2228001.

Our individual SNP analyses and overall gene analyses also indicated an association between *POLD1* and bladder cancer risk, which seem stronger among men. The polymerase (DNA directed), delta 1, catalytic subunit gene (*POLD*) (HGNC: 9175) is located on chromosome 19q13 and encodes the catalytic and proofreading subunit of Pol δ, which has polymerase and 3′-exonuclease activity.[34] We report associations with bladder cancer risk for two SNPs: rs2546551, an intronic SNP and rs2244095 SNP, which is 3′-downstream of *POLD*, within the Spi-B transcription factor (Spi-1/PU.1 related) gene (*SPIB*) (HGNC: 11242). Both SNPs are unlinked among Chinese and among NHW (HapMap CHBJPT $r^2 = 0.22$, $D' = 0.93$; CEU $r^2 = 0.12$, $D' = 1.00$). These SNPs are not linked with previously SNPs investigated in relation to bladder cancer risk, for which no associations were reported.[35–37]

We found that SNP rs7832529 in *POLB* associated with bladder cancer risk, mostly among Chinese. Summary estimates at the gene level using ARTP supported this finding. The polymerase (DNA directed) beta gene (*POLB*) (HGNC: 9174) is located on chromosome 8p11 and encodes a DNA polymerase involved in short patch and long patch BER.[38] Bladder cancer tumors and cell lines frequently encounter deletions in chromosomal region 8p, with 8p11-12 being one of the affected regions.[39] Located 3′-downstream from *POLB*, SNP rs7832529 is actually located within the solute carrier family 20 (phosphate transporter), member 2 gene (*SLC20A2*) (HGNC: 10947). To our knowledge, *SLC20A2* has not been linked with bladder cancer. Several other *POLB* SNPs have been reported to be associated with bladder cancer risk among Caucasians, but neither are in LD with rs7832529.[37,40] It remains to be determined whether rs7832529 is tagging a causal SNP in *POLB* or *SLC20A2*.

We also report that three *OGG1* SNPs (rs2072668, rs6809452 and rs1052133) were inversely associated with bladder cancer risk among Chinese females, with a stronger association for rs6809452. The 8-oxoG DNA glycosylase1 gene (*OGG1*) (HGNC: 12816) is located on chromosome 3p26.[41] The OGG1 protein participates in the removal of 8-oxoguanine (8-oxo-G) DNA damage that can result from ROS exposure. The intronic *OGG1* rs2072668 and rs6809452 SNPs were in strong LD with the non-synonymous and putative functional *OGG1* Ser326Cys SNP (rs1052133) (HapMap CHBJPT $r^2 = 0.98$, $D' = 1.00$ for rs2072668 and $r^2 = 0.88$, $D' = 1.00$ for rs6809452). SNP rs6809452, for which we found the strongest association, is actually an intronic SNP within the transcriptional adapter 3-like gene *(TADA3L)*. The *OGG1* Ser326Cys rs1052133 Cys allele has been reported to code for a protein with decreased ability to repair oxidative DNA damage.[42–46] A meta-analysis of various cancers reported that Ser326Cys was significantly associated with overall cancer risk and lung cancer risk but was not associated with bladder cancer risk.[47] Three epidemiological studies have reported associations between this SNP and bladder cancer risk among Caucasians, with stronger associations among smokers.[35–37]

Finally, we observed strong evidence that one SNP in the *XRCC6* gene, rs22284082, modified the effect of cigarette smoking. We found that among carriers of one or two copies of the C allele (major allele) there was a stronger and more significant association with tobacco smoking that among carriers of two copies of the T allele. The X-ray repair complementing defective repair in Chinese hamster cells 6 gene (*XRCC6*) (HGNC: 4055) is on chromosome region 22q13. SNP rs22284082 is located 3′-downstream from *XRCC6* and it maps to the sterol regulatory element binding transcription factor 2 gene (*SREBF2*) (HGNC: 11290), in intron 1. *SRBF2* encodes a transcription factor SREBP-2, a basic helix-loop-helix-leucine zipper protein that can stimulate transcription of sterol regulated genes and monitor lipid homeostasis.[48] In addition, SREBP-2 can also regulate autophagy related genes in times of nutrient depletion.[49] *SREBF2* has not been investigated in relation with bladder cancer; however, it has been reported to be involved in the loss of sterol feedback regulation in cancer cells.[50] It remains to be determined if the interaction with smoking we see for this SNP is capturing an effect of a causal SNP in *XRCC6* or *SREBF2*.

Our study had several strengths. Among them, was the use of two population-based case-control studies conducted in parallel in two world regions with contrasting bladder cancer incidence, using comparable instruments to assess

smoking exposure. Another one is the use of a comprehensive tagSNP approach that captured 85–100% genetic variation in genes that play key roles in four major DNA repair pathways, with appropriate consideration of multiple testing. Although we recognize that our tagSNP selection was done before the release of the 1,000 genomes project, which includes rare variants. Therefore, compared to this reference database, our overall genetic coverage would be lower. Finally, given that most studies on DNA repair susceptibility genes and bladder cancer have been conducting among NHW, our study contributes novel data about genetic risk factors among Chinese. Among the limitations of our study we include the fact that not all DNA repair genes from each pathway were captured, albeit all those that play essential roles were included and the fact that we were underpowered to explore higher order interactions between genes and exposures. Lastly, in spite of our approaches for multiple testing correction, we cannot discard the possibility that some of our findings might be false positives. Replication in other studies will help confirm our findings.

In conclusion, we found support that two regions that map close to or within BER genes (*POLB, OGG1*) and one region in an NER gene (*XPC*) are associated with bladder cancer risk, primarily among Chinese. Given that these associations were not modified by smoking, they suggest that there are other environmental factors that elicit the BER and NER pathways and might be relevant bladder cancer risk factors. We also find evidence that one SNP that tags both the *XRCC6* and *SREBF2* genes strongly modifies the association between bladder cancer risk and tobacco smoke. Given the role XRCC6 plays in the NHEJ pathway, this finding suggests that tobacco smoking may induce bladder cancer through the formation of double strand breaks. Further investigation in independent study populations will help confirm these findings and guide future studies to identify the causal variants responsible for these associations and all the relevant exposures that elicit the action of these DNA repair pathways.

Acknowledgements

The authors thank Ms. Susan Roberts for oversight of study enrollment, Ms. Peggy Wan for data analysis and management and Dr. Mimi C. Yu and Dr. Ronald K. Ross for their efforts in starting the LABC and SBC studies and feedback in the design of the current study. Roman Corral received support from NIH grant T32GM067587. Drs. Stern and Thomas received support from grant 5P30 ES07048 from NIEHS. Dr. Gago-Dominguez received support from FIS Intrasalud (PS09/02368) and the Botin Foundation. Drs. Thomas, Conti, Lewinger, and Cortessis received support from NIH grant 1R01 ES019876.

References

1. Ploeg M, Aben KK, Kiemeney LA. The present and future burden of urinary bladder cancer in the world. *World J Urol* 2009;27:289-93.
2. Liu L, Zhang J, Deapen D, *et al.* Cancer in Los Angeles county: incidence and mortality by race/ethnicity 1988-2000. University of Southern California, 2003.
3. Parkin DM, Whelan SL, Ferlay J, *et al.* Cancer Incidence in Five Continents, Volume VIII IARC Scientific Publications No. 155 Lyon: International Agency for Research on Cancer, 2002.
4. Lerner SP, Schoenberg MP, Sternberg CN. Textbook of bladder cancer. Abingdon, Oxon; Boca Raton: Taylor & Francis, 2006.
5. IARC Working Group on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, International Agency for Research on Cancer. Tobacco smoking. Lyon: World Health Organization, International Agency for Research on Cancer, 1986.
6. Freedman ND, Silverman DT, Hollenbeck AR, *et al.* Association between smoking and risk of bladder cancer among men and women. *JAMA* 2011;306:737-45.
7. Scelo G, Brennan P. The epidemiology of bladder and kidney cancer. *Nat Clin Pract Urol* 2007;4:205-17.
8. Gago-Dominguez M, Castelao JE, Yuan JM, *et al.* Use of permanent hair dyes and bladder-cancer risk. *Int J Cancer* 2001;91:575-9.
9. Vineis P, Talaska G, Malaveille C, *et al.* DNA adducts in urothelial cells: relationship with biomarkers of exposure to arylamines and polycyclic aromatic hydrocarbons from tobacco smoke. *Int J Cancer* 1996;65:314-6.
10. Pryor WA, Hales BJ, Premovic PI, *et al.* The radicals in cigarette tar: their nature and suggested physiological implications. *Science* 1983;220:425-7.
11. Maeda H, Sawa T, Yubisui T, *et al.* Free radical generation from heterocyclic amines by cytochrome b5 reductase in the presence of NADH. *Cancer Lett* 1999;143:117-21.
12. Burger MS, Torino JL, Swaminathan S. DNA damage in human transitional cell carcinoma cells after exposure to the proximate metabolite of the bladder carcinogen 4-aminobiphenyl. *Environ Mol Mutagen* 2001;38:1-11.
13. Lin J, Kadlubar FF, Spitz MR, *et al.* A modified host cell reactivation assay to measure DNA repair capacity for removing 4-aminobiphenyl adducts: a pilot study of bladder cancer. *Cancer Epidemiol Biomarkers Prev* 2005;14:1832-6.
14. Stern MC, Lin J, Figueroa JD, *et al.* Polymorphisms in DNA repair genes, smoking, and bladder cancer risk: findings from the international consortium of bladder cancer. *Cancer Res* 2009;69:6857-64.
15. Xing J, Dinney CP, Shete S, *et al.* Comprehensive pathway-based interrogation of genetic variations in the nucleotide excision DNA repair pathway and risk of bladder cancer. *Cancer* 2012;118:205-15.
16. Castelao JE, Yuan JM, Skipper PL, *et al.* Gender- and smoking-related bladder cancer risk. *J Natl Cancer Inst* 2001;93:538-45.
17. Tao L, Xiang YB, Wang R, *et al.* Environmental tobacco smoke in relation to bladder cancer risk-the Shanghai bladder cancer study [corrected]. *Cancer Epidemiol Biomarkers Prev* 2010;19:3087-95.
18. Yuan JM, Wang XL, Xiang YB, *et al.* Preserved foods in relation to risk of nasopharyngeal carcinoma in Shanghai, China. *Int J Cancer* 2000;85:358-63.
19. Edlund CK, Lee WH, Li D, *et al.* Snagger: a user-friendly program for incorporating additional information for tagSNP selection. *BMC Bioinformatics* 2008;9:174.
20. Oliphant A, Barker DL, Stuelpnagel JR, *et al.* BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques* 2002;Suppl:56-8, 60-1.
21. Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. *Am J Hum Genet* 2007;81:1158-68.
22. Yu K, Li Q, Bergen AW, Pfeiffer RM, *et al.* Pathway analysis by adaptive combination of P-values. *Genet Epidemiol* 2009;33:700-9.
23. Tang W, Fu YP, Figueroa JD, *et al.* Mapping of the UGT1A locus identifies an uncommon coding variant that affects mRNA expression and protects from bladder cancer. *Hum Mol Genet* 2012;21:1918-30.
24. Rothman N, Garcia-Closas M, Chatterjee N, *et al.* A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. *Nat Genet* 2010;42:978-84.
25. Rafnar T, Vermeulen SH, Sulem P, *et al.* European genome-wide association study identifies SLC14A1 as a new urinary bladder cancer susceptibility gene. *Hum Mol Genet* 2011;20:4268-81.

26. Menashe I, Figueroa JD, Garcia-Closas M, *et al.* Large-scale pathway-based analysis of bladder cancer genome-wide association data from five studies of European background. *PloS one* 2012;7: e29396.
27. Kiemeney LA, Thorlacius S, Sulem P, *et al.* Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. *Nat Genet* 2008;40:1307-12.
28. Kiemeney LA, Sulem P, Besenbacher S, *et al.* A sequence variant at 4p16.3 confers susceptibility to urinary bladder cancer. *Nat Genet* 2010;42: 415-9.
29. Garcia-Closas M, Ye Y, Rothman N, *et al.* A genome-wide association study of bladder cancer identifies a new susceptibility locus within SLC14A1, a urea transporter gene on chromosome 18q12.3. *Hum Mol Genet* 2011;20: 4282-9.
30. Friedberg EC. How nucleotide excision repair protects against cancer. *Nat Rev Cancer* 2001;1: 22-33.
31. Sugasawa K, Ng JM, Masutani C, *et al.* Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell* 1998;2:223-32.
32. Araki M, Masutani C, Takemura M, *et al.* Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *J Biol Chem* 2001;276:18665-72.
33. Qiao B, Scott GB, Elliott F, *et al.* Functional assays to determine the significance of two common XPC 3'UTR variants found in bladder cancer patients. *BMC Med Genet* 2011; 12:84.
34. Burgers PM. Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem* 2009;284: 4041-5.
35. 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;78:464-79.
36. Huang M, Dinney CP, Lin X, *et al.* High-order interactions among genetic variants in DNA base excision repair pathway genes and smoking in bladder cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 2007;16:84-91.
37. Figueroa JD, Malats N, Real FX, *et al.* Genetic variation in the base excision repair pathway and bladder cancer risk. *Hum Genet* 2007;121: 233-42.
38. Dogliotti E, Fortini P, Pascucci B, *et al.* The mechanism of switching among multiple BER pathways. *Prog Nucleic Acid Res Mol Biol* 2001; 68:3-27.
39. Wagner U, Bubendorf L, Gasser TC, *et al.* Chromosome 8p deletions are associated with invasive tumor growth in urinary bladder cancer. *Am J Pathol* 1997;151:753-9.
40. Michiels S, Laplanche A, Boulet T, *et al.* Genetic polymorphisms in 85 DNA repair genes and bladder cancer risk. *Carcinogenesis* 2009;30:763-8.
41. Boiteux S, Radicella JP. The human OGG1 gene: structure, functions, and its implication in the process of carcinogenesis. *Arch Biochem Biophys* 2000;377:1-8.
42. Lee AJ, Hodges NJ, Chipman JK. Interindividual variability in response to sodium dichromate-induced oxidative DNA damage: role of the Ser326Cys polymorphism in the DNA-repair protein of 8-oxo-7,8-dihydro-2′-deoxyguanosine DNA glycosylase 1. *Cancer Epidemiol Biomarkers Prev* 2005;14:497-505.
43. Yamane A, Kohno T, Ito K, *et al.* Differential ability of polymorphic OGG1 proteins to suppress mutagenesis induced by 8-hydroxyguanine in human cell in vivo. *Carcinogenesis* 2004;25: 1689-94.
44. Hill JW, Evans MK. Dimerization and opposite base-dependent catalytic impairment of polymorphic S326C OGG1 glycosylase. *Nucleic Acids Res* 2006;34:1620-32.
45. Sidorenko VS, Grollman AP, Jaruga P, *et al.* Substrate specificity and excision kinetics of natural polymorphic variants and phosphomimetic mutants of human 8-oxoguanine-DNA glycosylase. *Febs J* 2009;276:5149-62.
46. Kershaw RM, Hodges NJ. Repair of oxidative DNA damage is delayed in the Ser326Cys polymorphic variant of the base excision repair protein OGG1. *Mutagenesis* 2012;27:501-10.
47. Wei B, Zhou Y, Xu Z, *et al.* The effect of hOGG1 Ser326Cys polymorphism on cancer risk: evidence from a meta-analysis. *PLoS One* 2011;6: e27545.
48. Horton JD, Goldstein JL, Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest* 2002;109:1125-31.
49. Seo YK, Jeon TI, Chong HK, *et al.* Genome-wide localization of SREBP-2 in hepatic chromatin predicts a role in autophagy. *Cell Metab* 2011;13: 367-75.
50. Chen Y, Hughes-Fulford M. Human prostate cancer cells lack feedback regulation of low-density lipoprotein receptor and its regulator, SREBP2. *Int J Cancer* 2001;91:41-5.