Genome-wide association study identified SNP on 15q24 associated with bladder cancer risk in Japanese population

Koichi Matsuda1, Atsushi Takahashi2, Candace D. Middlebrooks3, Wataru Obara4, Yasutomo Nasu5, Keiji Inoue6, Kenji Tamura6, Ichiro Yamasaki6, Yoshio Naya7, Chizu Tanikawa1, Ri Cui1, Jonine D. Figueroa3, Debra T. Silverman3, Nathaniel Rothman3, Mikio Namiki8, Yoshihiko Tomita9, Hiroyuki Nishiyama10, Kenjiro Kohri11, Takashi Deguchi12, Masayuki Nakagawa13, Masayoshi Yokoyama14, Tsuneharu Miki7, Hiromi Kumon5, Tomoaki Fujioka4, Ludmila Prokunina-Olsson3, Michiaki Kubo2, Yusuke Nakamura15,16,* and Taro Shuin6

1Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, 2Center for Integrative Medical Science, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan, 3Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, USA, 4Department of Urology, Iwate Medical University School of Medicine, Iwate, Japan, 5Department of Urology, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan, 6Department of Urology, School of Medicine, Kochi University, Kochi, Japan, 7Department of Urology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan, 8Department of Integrative Cancer Therapy and Urology, Kanazawa University Graduate School of Medical Science, Ishikawa, Japan, 9Department of Urology, Yamagata University Faculty of Medicine, Yamagata, Japan, 10Department of Urology, Faculty of Medicine, University of Tsukuba, Ibaragi, Japan, 11Department of Nephro-urolgy, Nagoya City University Graduate School of Medical Sciences, Aichi, Japan, 12Department of Urology, Graduate School of Medicine, Gifu University, Gifu, Japan, 13Department of Urology, Graduate School of Medical and Dental Sciences, Kagoshima University, Kagoshima, Japan, 14Department of Urology, Ehime University Graduate School of Medicine, Ehime, Japan, 15Department of Medicine and 16Department of Surgery, Center for Personalized Therapeutics, The University of Chicago, Chicago, IL, USA

Received June 30, 2014; Revised September 18, 2014; Accepted September 29, 2014

Through genome-wide association analysis and an independent replication study using a total of 1131 bladder cancer cases and 12558 non-cancer controls of Japanese populations, we identified a susceptibility locus on chromosome 15q24. SNP rs11543198 was associated with bladder cancer risk with odds ratio (OR) of 1.41 and P-value of 4.03 × 10−9. Subgroup analysis revealed rs11543198 to have a stronger effect in male smokers with OR of 1.66. SNP rs8041357, which is in complete linkage disequilibrium (r² = 1) with rs11543198, was also associated with bladder cancer risk in Europeans (P = 0.045 for an additive and P = 0.025 for a recessive model), despite much lower minor allele frequency in Europeans (3.7%) compared with the Japanese (22.2%). Imputational analysis in this region suggested CYP1A2, which metabolizes tobacco-derived carcinogen, as a causative candidate gene. We also confirmed the association of previously reported loci, namely SLC14A1, APOBEC3A, PSCA and MYC, with bladder cancer. Our finding implies the crucial roles of genetic variations on the chemically associated development of bladder cancer.

*To whom correspondence should be addressed at: Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Knapp Center for Biomedical Discovery, Room 6130, MC2115, 900 E. 57th Street, Chicago, IL 60637, USA. Tel: +1 7738341405; Fax: +1 7737029268; Email: ynakamura@bsd.uchicago.edu

© The Author 2014. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
INTRODUCTION

Bladder cancer is one of the most frequent cancers, which causes \sim 150,000 death per year in the world (1). Both environmental and genetic factors are involved in the development of bladder cancer, and tobacco smoking is known to be the most important factor to increase the risk of bladder cancer; current or former smokers have a 2- to 6-fold higher risk than never-smokers (2,3). In addition, occupational exposures to industrial chemicals (4–6), arsenic contamination in drinking water (7) and infectious diseases (8) also increase the bladder cancer risk. Bladder cancer incidence in males is nearly 3-fold higher than that in females (3), probably due to the higher prevalence of tobacco smoking and occupational exposure in males. On the other hand, familial aggregation of bladder cancer has also been reported (9,10), suggesting the importance of genetic factors in bladder cancer development.

NAT2 and GSTM1 are involved in the detoxification of carcinogen (11), and an NAT2 slow-metabolizer genotype and a GSTM1 null genotype were indicated their association with an increased risk of bladder cancer (12,13). In addition, recent genome-wide association studies (GWASs) in European populations have identified multiple genetic factors associated with bladder cancer (14–18). However, no GWAS has yet been conducted in Asian populations. So, to identify genetic factors associated with the risk of bladder cancer, we conducted a genome-wide association study in the Japanese population.

RESULTS

In this study, we performed a two-stage GWAS approach. A study design and the information of subjects used in this analysis are shown in Supplementary Material, Fig. S1 and Table S1. Case samples in the screening and replication stages were recruited from a collaborative network consisting of 11 university-affiliated hospitals in Japan. The genotyping information for non-cancer controls was obtained from the BioBank Japan project. In the screening stage, a total of 539 bladder cancer cases were genotyped by using Human OmniExpress Exome chip. After an initial standard quality control procedure, we obtained genotyping results of 554,389 SNP loci in 531 cases and compared them with those of 5581 controls. Principle component analysis indicated that all subjects used in GWAS were of the Asian ancestry (Supplementary Material, Fig. S2). We performed statistical analysis using a Cochran–Armitage trend test and obtained the genomic inflation factor lambda to be 1.0493, indicating the low possibility of population stratification (Supplementary Material, Fig. S3). From the association analysis, 82 SNPs in 45 distinct genomic regions were indicated suggestive associations ($P \leq 5 \times 10^{-5}$). Among them, three SNPs on chromosome 6p21 passed the genome-wide significance threshold ($P = 1.76\text{–}3.39 \times 10^{-8}$, Fig. 1 and Supplementary Material, Table S2).

To validate these associations, we selected 64 SNPs through linkage disequilibrium analysis with the criteria of pair-wise $r^2$ of \sim 0.8. In the replication analysis, we attempted genotyping of a new set of 592 cases and 6964 controls for these 64 SNPs and successfully obtained the genotype information at 59 SNP loci. Among these 59 SNPs, the only one SNP, rs11543198, on chromosome 15q24 revealed significant association with a $P$-value of $1.22 \times 10^{-4}$ ($P = 8.47 \times 10^{-4}$; $0.05/59$, Table 1 and Supplementary Material, Table S3). The combined analysis of the GWAS and replication data indicated SNP rs11543198 to have the genome-wide significance of the association ($P_{\text{meta}} = 4.03 \times 10^{-9}$, odds ratio (OR) = 1.41, Table 1).

Figure 1. Manhattan plot showing the genome-wide $P$-values of association. The $P$-values were calculated by Cochran–Armitage trend test. The $x$-axis represents the $-\log_{10} P$-values of 554,389 SNPs, and $x$-axis shows their chromosomal positions. The horizontal blue line shows the threshold of $P \leq 5 \times 10^{-5}$ for selecting 84 SNPs for replication analysis.
For validation, we examined the association of this locus with bladder cancer risk in 3508 bladder cancer cases and 5101 controls of European ancestry (17). As rs11543198 was not genotyped in the European GWAS, we used SNP rs8041357, which is in complete linkage disequilibrium ($D^2 = 1, r^2 = 1$) with rs11543198 in Asians and Europeans based on the 1000 Genomes Project data (19). In agreement with the 1000 Genomes data, the frequency of the minor allele $G$ was 22.2% in the Japanese controls, compared with 3.7% in Europeans (Supplementary Material, Table S4).

Owing to much lower allele frequency, the association in Asians was weaker but the direction and magnitude of association were similar in both populations.

As gender and smoking status are known to be important risk factors for bladder cancer, we conducted subgroup analysis. As a result, SNP rs11543198 showed a stronger effect among males and smokers (OR of 1.66 for smoker and 1.49 for never-smoker) than females (OR of 1.55, Supplementary Material, Table S5). Moreover, an A allele of SNP rs2069514 (CYP1A2*1C), which is known to have the lower CYP1A2 activity (20), was indicated to reduce the risk of tobacco-related cancers. Therefore, our result suggests that CYP1A2 on chromosome 15q24 may be a causative gene candidate to increase the risk of bladder cancer.

In addition, we also examined previously reported susceptible loci in our GWAS cohort (14–18). We successfully genotyped 10 SNPs as shown in Table 3 and confirmed the significant association of the loci in the $SLC14A1$, $APOBEC3A$, $PSCA$ and $MYC$ genes with P-values of $<0.05$ (Table 3). Two SNPs on $TACC3$ and $UGT1A$ loci revealed weak trends, but the remaining four loci were not replicated. These results may imply the genetic heterogeneity between European and Asian populations.

### Table 1. Summary of GWAS and replication analyses

<table>
<thead>
<tr>
<th>SNP</th>
<th>Case</th>
<th>Control</th>
<th>$P_{\text{trend}}$</th>
<th>ORb (95% CI)</th>
<th>$P_{\text{het}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11543198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GWAS</td>
<td>16</td>
<td>140</td>
<td>3.24</td>
<td>1.48</td>
<td>(1.25–1.75)</td>
</tr>
<tr>
<td>Replication</td>
<td>21</td>
<td>158</td>
<td>4.03</td>
<td>1.41</td>
<td>(1.26–1.59)</td>
</tr>
</tbody>
</table>

A total of 1131 bladder cancer cases and 12 545 controls were analyzed.  

$^a$Odds ratios (OR) and confidence interval (CI) are calculated using the non-susceptible allele (A) as reference.  

$^b$The $P$-values of heterogeneities across four subgroups was examined by using the Breslow–Day test.  

$^c$The $P$-values of heterogeneities across two stages were examined by using the Breslow–Day test.  

### Table 2. Subgroup analysis stratified by gender and smoking status

<table>
<thead>
<tr>
<th>SNP</th>
<th>Case</th>
<th>Control</th>
<th>$P_{\text{trend}}^a$</th>
<th>ORb (95% CI)</th>
<th>$P_{\text{het}}^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11543198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16</td>
<td>148</td>
<td>3.01 × 10$^{-8}$</td>
<td>1.58 (1.34–1.87)</td>
<td>0.07</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>45</td>
<td>0.32</td>
<td>1.16 (0.86–1.55)</td>
<td>0.10</td>
</tr>
<tr>
<td>Smoker</td>
<td>11</td>
<td>95</td>
<td>8.63 × 10$^{-7}$</td>
<td>1.65 (1.34–1.87)</td>
<td>0.15</td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>11</td>
<td>98</td>
<td>0.01</td>
<td>1.30 (0.86–1.55)</td>
<td>0.04</td>
</tr>
<tr>
<td>Male/smoker</td>
<td>10</td>
<td>90</td>
<td>1.63 × 10$^{-6}$</td>
<td>1.66 (1.35–2.04)</td>
<td>0.19</td>
</tr>
<tr>
<td>Male/nonsmoker</td>
<td>6</td>
<td>58</td>
<td>3.24 × 10$^{-3}$</td>
<td>1.49 (1.14–1.95)</td>
<td>0.30</td>
</tr>
<tr>
<td>Female/smoker</td>
<td>1</td>
<td>5</td>
<td>0.54</td>
<td>1.29 (0.57–2.95)</td>
<td>0.20</td>
</tr>
<tr>
<td>Female/nonsmoker</td>
<td>5</td>
<td>40</td>
<td>0.47</td>
<td>1.12 (0.82–1.54)</td>
<td>0.14</td>
</tr>
</tbody>
</table>

A total of 739 bladder cancer cases and 11 380 controls were analyzed.  

$^a$P-value obtained from Cochrane–Armitage trend test.  

$^b$Odds and CI are calculated using the non-susceptible allele (A) as reference.  

$^c$The $P$-value of heterogeneities across four subgroups was examined by using the Breslow–Day test.
DISCUSSION

Here, we reported the result of GWAS analysis using a total of 1131 Japanese bladder cancer cases and 12,558 controls. Our data indicated that SNP rs11543198 on chromosome 15q24 was significantly associated with bladder cancer risk. SNP rs11543198 is located within the \textit{CLK3} gene, which encodes a serine/threonine protein kinase. CLK3 regulates localization of SR family of splicing factors; however, its association with human carcinogenesis was not reported so far. Although SNP rs11543198 alters amino acid sequence of minor isoforms of CLK3 (X1, X2, X5 and X6), this SNP does not affect amino acid sequence of catalytic domain. Therefore, SNP rs11543198 is not likely to be a causative variation.

Imputation analysis of 29 SNPs in the 341-kb region including this SNP locus suggested an SNP near the \textit{CYP1A2} gene to be a causative candidate because of its stronger association than the marker SNP rs11543198 as well as its biological relevance. Interestingly, \textit{CYP1A2} belongs to the cytochrome P450 superfamily, members of which encode monooxygenases involved in metabolism of various substrates including drugs and play essential roles in the synthesis of cholesterol, steroids and other lipids (23,24). CYP1A2, which is highly expressed in the liver, is activated by the exposure to PAHs (25,26) and metabolizes them. It is well known that CYP1A2 metabolizes heterocyclic aromatic amines contained in tobacco smoke (27,28) and generates carcinogenic intermediates (21,22). In addition, genetic variations in the \textit{CYP1A2} locus are known to affect its enzymatic activity (29,30). CYP1A2 activity exhibits a significant degree of inter-individual diversity owing to both environmental and genetic factors. Hence, associations of \textit{CYP1A2} variations with various cancers including bladder, breast, colorectal and lung cancers have been repeatedly investigated, but the results are controversial (31–34). Our first GWAS in the Asian population revealed that \textit{CYP1A2} locus is significantly associated with bladder cancer. Moreover, SNP rs11543198 showed the stronger effect on smokers compared with never-smokers in both males and females. Considering the role of CYP1A2 in the metabolism of tobacco-derived carcinogen, our finding suggested the interesting gene-environmental interaction on the development of bladder cancer.

Among genes associated with bladder cancer in previously reported GWAS, our analyses revealed that SNPs in the \textit{SLC14A1}, \textit{APOBEC3A-CBX6}, \textit{PSCA} and \textit{MYC} loci were associated with bladder cancer risk in the Japanese population (\textit{P} < 0.05). Association of these four SNPs with bladder cancer was also reported in the Chinese population (35,36). Thus, these variations are common bladder cancer loci among the European and Asian populations.

\textit{SLC14A1} functions as a urea transporter and regulates urine concentration and body-fluid balance. Although the function of SNP rs17674580 in SLC14A1 was not fully elucidated, our previous analysis revealed that \textit{SLC14A1-SLC14A2} locus was associated with blood urea nitrogen level (37). Thus, this genetic variation
variation would be associated with the kidney function and urine concentration and subsequently affect on the bladder cancer risk. SNP rs1014971 on chromosome 22q13.1 is located \( \approx 25 \text{ kb} \) centromeric to \textit{APOBEC3A}. As \textit{APOBEC3A} was expressed only in peripheral blood leukocytes and spleen (38), the role of this variation in the carcinogenesis of bladder cancer is not yet clarified.

SNPs rs2294008 and rs9642880 are located on chromosome 8q24, and these loci are not in the same linkage disequilibrium (LD) block. SNP rs9642880, which is located in the LD block adjacent to \textit{MYC}, a well-known oncogene, was shown to be associated with \textit{MYC} mRNA and protein expression (39). SNP rs2294008 is associated with cell surface localization and higher expression level of PSCA (40,41). PSCA, which is up-regulated in various tumors including bladder cancer, was shown to be involved in cell renewal and proliferation (42). Therefore, the association of these SNPs with bladder cancer risk can be explained by the growth promoting effect of \textit{MYC} and cell surface PSCA.

Taken together, we here demonstrated the important roles of genetic factors on the development of bladder cancer. Particularly, the stronger effect of rs11543198 on smokers is very important because it implies a possibility that a small change in the life style, quitting or avoiding smoking, may contribute to the improvement of individuals carrying the risk genotype. Although further prospective analysis is necessary, we hope that our finding would further emphasize the significance of tobacco control.

### MATERIALS AND METHODS

**Study population**

Characteristics of each cohort are shown in Supplementary Material, Table S1. In this study, we conducted GWAS and replication analyses using a total of 1131 bladder cancer cases and 12,558 controls. Case samples in GWAS and replication were obtained from a collaboration network consisting of Iwate Medical University, Okayama University, Kochi Medical School, Kyoto Prefectural University of Medicine, Kanazawa University, Yamagata University, University of Tsukuba, Nagoya City University, Gifu University, Kagoshima University and Ehime University. Control samples in GWAS and replication consisted of healthy volunteers (\( n = 1919 \)) and subjects with other diseases (\( n = 10,639 \)), cerebral aneurysm, chronic obstructive pulmonary disease, glaucoma, nephrolithiasis, nephrotic syndrome, epilepsy, atopic dermatitis and Grave’s disease) obtained from Biobank Japan Project supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan. In the BioBank Japan Project, DNA and serum of patients were collected through a collaborating network of 66 hospitals throughout Japan. More than 200,000 individuals with 47 common diseases, irrespective of prior treatment, were enrolled in this project from 2003. A list of participating hospitals is provided at the BioBank Japan website (http://www.biobankjp.org/english/index.html). Subjects with a history of any cancers were excluded from controls. Smoking information of both cases and controls was obtained by oral interview. This project was approved by the ethical committees at each institute.
SNP genotyping
Genomic DNA was extracted from peripheral blood leukocytes or normal tissue using a standard method. In GWAS, we genotyped 539 bladder cancer cases and 5769 non-cancer controls (cerebral aneurysm, primary sclerosing cholangitis, chronic obstructive pulmonary disease, glaucoma and healthy volunteers) using Human OmniExpress and HumanExome (Supplementary Material, Fig. S1). A total of 951 117 SNPs including 925 436 autosomal SNPs were genotyped by both platforms. Among the 925 436 autosomal SNPs, 223 273 SNPs were monomorphic in our case–control sample set. We excluded the following samples from analysis: closely related samples, gender mismatch and subjects whose ancestries were estimated to be distinct from East-Asian populations using principal component analysis. Then, we applied SNP quality control as follows: call rate ≥ 0.99 in case and control samples, minor allele frequency (MAF) of <0.01 and P-value of Hardy–Weinberg Equilibrium in control group ≥ 1 × 10⁻⁶. Consequently, 554 389 SNPs on autosomal chromosomes passed the quality control filters. Among 82 SNPs showing P < 5 × 10⁻⁵, we selected 64 SNPs by linkage disequilibrium analysis with the criteria of pair-wise r² of >0.8. In the replication analysis, we genotyped 64 SNPs in 592 bladder cancer cases by using multiplex PCR-based Invader assay (Third Wave Technologies). A total of 6964 non-cancer controls (nephrolithiasis, epilepsy, atrophic dermatitis and Grave’s disease) from Biobank Japan were genotyped by Human OmniExpress exome beadchip.

Statistical analysis
The association of SNPs with bladder cancer risk was tested by Cochran–Armitage trend test. To characterize population structure in the GWAS cohort, we performed principal component analysis (43). In the GWAS, the genetic inflation factor λ was derived by P-values obtained by Cochran–Armitage trend test for all the tested SNPs. The quantile–quantile plot was drawn using R program. The ORs were calculated using the non-parametric method. In GWAS, the genetic inflation factor was examined by using the Breslow–Day test. The combined analysis of GWAS and replication stage was verified by conducting the Mantel–Haenszel method. Heterogeneity across two stages was examined by using the Breslow–Day test (44). We considered P = 5 × 10⁻⁵ (GWAS and meta-analysis) and 8.33 × 10⁻⁴ (0.05/59, replication analysis) as the significant threshold after Bonferroni correction for multiple testing.

Replication analysis in Europeans
Genotypes of rs8041357 in 3508 bladder cancer cases and 5101 controls of European ancestry were extracted from the previous GWAS data set (17). Each participating study obtained informed consent from study participants and approval from its respective Institutional Review Board for this study. Genome-wide genotyping was conducted using HumanHap 1M or HumanHap610-Quad BeadCHIP (Supplementary Material, Table S1). Genotypes of rs8041357 were in Hardy–Weinberg equilibrium in controls (P > 0.05). Association was evaluated for an additive model adjusting for age, sex, smoking status (ever/never), study sites and significant EVs as in the original GWAS (17). Owing to low allele frequency for allele G (3.7%), we also analyzed the results using a recessive genetic model, evaluating the effect in risk homozygotes (AA) versus a combined group of rare homozygotes (GG) and heterozygotes (AG), adjusting for the same covariates.

Imputation analysis
Imputation of ungenotyped SNPs was conducted by MACH (45) and minimac (46) using data of JPT/CHS/CHD subjects from 1000 genome project phase 1 (release 16 March 2012) as a reference. We excluded SNPs that met the following criteria: MAF < 0.01, Hardy–Weinberg Equilibrium P-value < 1 × 10⁻⁶ or large allele frequency difference between reference panel and GWAS (> 0.16).

Software
For general statistical analysis, we employed R statistical environment version 2.9.1 (cran.r-project.org). The Haploview software version 4.2 (47) was used to draw Manhattan plot. Primer3-webv3.0 (http://frodo.wi.mit.edu) web tool was used to design primers. MACH (45) (http://www.sph.umich.edu/csg/abecasis/MACH/), minimac (46) (http://genome.sph.umich.edu/wiki/Minimac) and mach2dat (http://genome.sph.umich.edu/wiki/Mach2dat:_Association_with_MACH_output) were used for imputation analysis.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
We thank all the patients and the members of the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan, who donated their DNA for this work. We also thank Satomi Takahashi (the University of Tokyo) and the technical staff of the Laboratory for Genotyping Development, Center for Integrative Medical Science, RIKEN for their technical support.

Conflict of Interest statement. None declared.

FUNDING
This work was conducted as a part of the BioBank Japan Project that was supported by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government. The work was supported in part by the Intramural Research Program (IRP) of the National Cancer Institute of the US National Institutes of Health.

REFERENCES


