Common genetic variants in the *PSCA* gene influence gene expression and bladder cancer risk

Yi-Ping Fu[a], Indu Kohaar[a], Nathaniel Rothman[b], Julie Earl[c], Jonine D. Figueroa[b], Yuanqing Ye[d], Núria Malats[e], Wei Tang[a], Luyang Liu[a], Montserrat Garcia-Closas[b,f], Brian Muchmore[a], Nilanjan Chatterjee[b], McAnthony Tarway[a], Manolis Kogevinas[g,h,i,j], Patricia Porter-Gill[a], Dalsu Baris[b], Adam Mumy[a], Demetrius Albanes[b], Mark P. Purdue[b], Amy Hutchinson[k], Alfredo Carrato[l], Adonina Tardón[h,m], Consol Serra[n], Reina García-Closas[o], Josep Lloreta[p], Alison Johnson[q], Molly Schwenn[r], Margaret R. Karagas[s], Alan Schned[s], W. Ryan Diver[t], Susan M. Gapstur[t], Michael J. Thun[t], Jarmo Virtamo[u], Stephen J. Chanock[a], Joseph F. Fraumeni, Jr.[b,1], Debra T. Silverman[b], Xifeng Wu[d], Francisco X. Real[c,n], and Ludmila Prokunina-Olsson[a,1]

[a]Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, and [b]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892; [c]Epithelial Carcinogenesis Group, Molecular Pathology Programme, Centro Nacional de Investigaciones Oncológicas, Madrid 28029, Spain; [d]Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, TX 77030; [e]Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Center, Madrid 28029, Spain; [f]Division of Genetics and Epidemiology, Institute of Cancer Research, London SW7 3RP, United Kingdom; [g]Centre for Research in Environmental Epidemiology, Barcelona 08003, Spain; [h]Municipal Institute of Medical Research, Barcelona 08003, Spain; [i]Centro de Investiación de Biomédica en Red Epidemiología y Salud Pública, Barcelona 08003, Spain; [j]National School of Public Health, Athens 11521, Greece; [k]Core Genotyping Facility, SAIC-Frederick, Inc., National Cancer Institute at Frederick, Frederick, MD 21702; [l]Ramón y Cajal University Hospital, Madrid 28034, Spain; [m]Instituto Universitario de Oncología, Universidad de Oviedo, Oviedo 33003, Spain; [n]Departament de Ciències Experimentals i de la Salut, and [p]Hospital del Mar-Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra, Barcelona 08003, Spain; [o]Unidad de Investigación, Hospital Universitario de Canarias, La Laguna 38320, Spain; [q]Vermont Cancer Registry, Burlington, VT 05401; [r]Maine Cancer Registry, Augusta, ME 04333; [s]Dartmouth Medical School, Hanover, NH 03755; [t]Epidemiology Research Program, American Cancer Society, Atlanta, GA 30303; and [u]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki 00271, Finland

Contributed by Joseph F. Fraumeni, Jr., February 15, 2012 (sent for review January 11, 2012)

Genome-wide association studies have identified a SNP, rs2294008, on 8q24.3 within the prostate stem cell antigen (*PSCA*) gene, as a risk factor for bladder cancer. To fine-map this region, we imputed 642 SNPs within 100 Kb of rs2294008 in addition to 33 markers genotyped in one of the reported genome-wide association study in 8,652 subjects. A multivariable logistic regression model adjusted for rs2294008 revealed a unique signal, rs2978974 (r^2 = 0.02, D′ = 0.19 with rs2294008). In the combined analysis of 5,393 cases and 7,324 controls, we detected a per-allele odds ratio (OR) = 1.11 [95% confidence interval (CI) = 1.06–1.17, $P = 5.8 \times 10^{-5}$] for rs2294008 and OR = 1.07 (95% CI = 1.02–1.13, $P = 9.7 \times 10^{-3}$) for rs2978974. The effect was stronger in carriers of both risk variants (OR = 1.24, 95% CI = 1.08–1.41, $P = 1.8 \times 10^{-3}$) and there was a significant multiplicative interaction (P = 0.035) between these two SNPs, which requires replication in future studies. The T risk allele of rs2294008 was associated with increased *PSCA* mRNA expression in two sets of bladder tumor samples (n = 36, P = 0.0007 and n = 34, P = 0.0054) and in normal bladder samples (n = 35, P = 0.0155), but rs2978974 was not associated with *PSCA* expression. SNP rs2978974 is located 10 Kb upstream of rs2294008, within an alternative untranslated first exon of *PSCA*. The non-risk allele G of rs2978974 showed strong interaction with nuclear proteins from five cell lines tested, implying a regulatory function. In conclusion, a joint effect of two *PSCA* SNPs, rs2294008 and rs2978974, suggests that both variants may be important for bladder cancer susceptibility, possibly through different mechanisms that influence the control of mRNA expression and interaction with regulatory factors.

Urinary bladder cancer ranks as the ninth most common malignancy worldwide (1) and the sixth in the United States, with 73,510 new cases and 14,880 deaths expected in the United States in 2012 (2). Because of a 70% 10-y survival rate and 50–70% recurrence rate (3), patients require life-long surveillance and treatment, making bladder cancer one of the most expensive cancers to live with and a major economic burden on the health care system (4, 5). Cigarette smoking and occupational exposures to aromatic amines are established risk factors for bladder cancer (6).

Genetic factors are estimated to explain 7–31% of bladder cancer susceptibility (7, 8). Traditional candidate gene association studies focused on genes involved in detoxification of environmental carcinogens and DNA repair pathways (9–12). Recently, genome-wide association studies (GWAS) have advanced our knowledge of the genetic architecture of bladder cancer by discovering a number of new bladder cancer-associated loci (13–17).

One of these variants to achieve genome-wide significance with bladder cancer was SNP rs2294008 (C/T) within the prostate stem cell antigen (*PSCA*) gene on 8q24.3 (14, 15). PSCA is a glycosylphosphatidylinositol-anchored cell membrane glycoprotein initially identified as a prostate-specific cell-surface antigen (18), but later found to be expressed in many human tissues. Humanized monoclonal anti-PSCA antibodies are currently under evaluation in clinical trials for the treatment of pancreatic and prostate cancers (19, 20). Although many studies have supported the important role of *PSCA* in carcinogenesis, little is known about the function of *PSCA* in normal and tumor tissues.

In this study, we aimed to comprehensively explore the genetic landscape of the *PSCA* region to map genetic variants contributing to bladder cancer susceptibility, as well as to investigate the molecular phenotype of these variants and their role in cancer biology.

Results

Fine-Mapping and Genetic Association of the *PSCA* Region. *PSCA* is located in 8q24.3 locus, 12.5 Mb telomeric from the region associated with multiple cancers (21). Based on the combined reference panel of the 1000 Genomes Project and HapMap 3 CEU (Utah residents with Northern and Western European ancestry) data, we expanded the SNP coverage of the 200 Kb *PSCA* region (chr8: 143,658,933–143,858,933, University of California at Santa Cruz (UCSC) genome build hg18), which covers eight genes: *ARC*, *JRK*, *PSCA*, *LY6K*, *C8orf55*, *SLURP1*, *LYPD2*, and *LYNX1*. In addition

Author contributions: Y.-P.F., I.K., N.R., J.E., J.D.F., S.J.C., J.F.F., D.T.S., X.W., F.X.R., and L.P.-O. designed research; Y.-P.F., I.K., J.E., Y.Y., N.M., W.T., L.L., B.M., M.T., M.K., P.P.-G., D.B., A.M., D.A., M.P.P., A.H., A.C., A.T., C.S., R.G.-C., J.L., A.J., M.S., M.R.K., A.S., W.R.D., S.M.G., M.J.T., J.V., X.W., F.X.R., and L.P.-O. performed research; J.D.F., M.G.-C., and N.C. contributed new reagents/analytic tools; Y.-P.F., I.K., J.E., J.D.F., Y.Y., W.T., M.G.-C., N.C., X.W., F.X.R., and L.P.-O. analyzed data; N.R., Y.Y., N.M., M.G.-C., M.K., D.B., D.A., M.P.P., A.H., A.C., A.T., C.S., R.G.-C., J.L., A.J., M.S., M.R.K., A.S., W.R.D., S.M.G., M.J.T., J.V., S.J.C., J.F.F., D.T.S., X.W., and F.X.R. subject enrollment and data collection; and Y.-P.F., I.K., N.R., S.J.C., J.F.F., D.T.S., X.W., F.X.R., and L.P.-O. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: fraumeni@nih.gov or prokuninal@mail.nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1202189109/-/DCSupplemental.

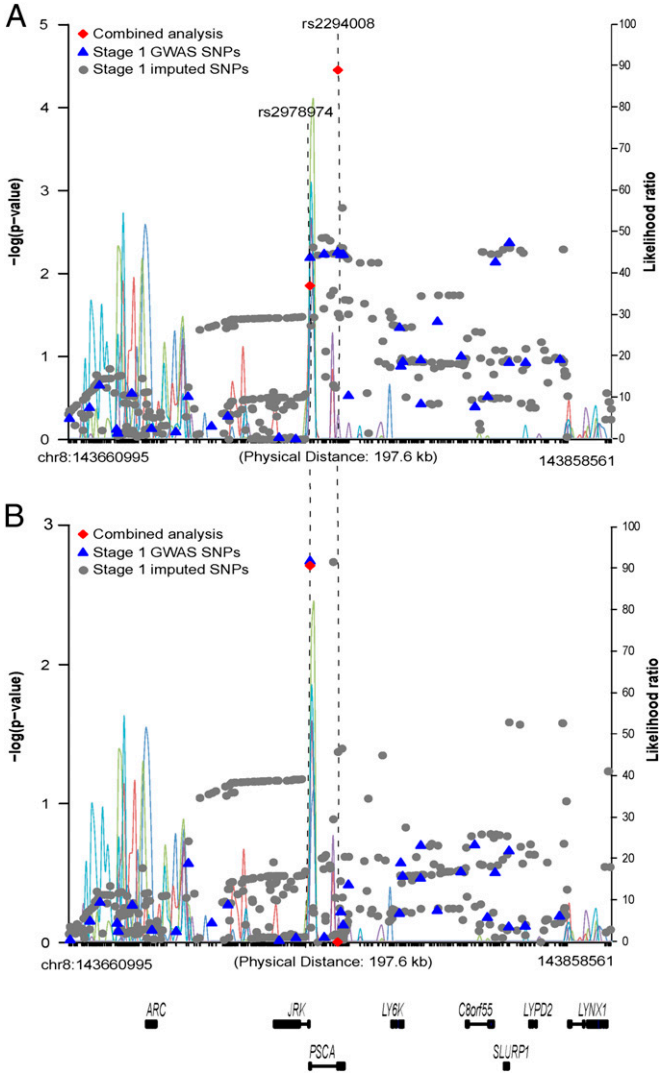

Fig. 1. Association results for bladder cancer risk and recombination plots of the *PSCA* region. Results for 33 GWAS-genotyped markers (blue triangles), 413 imputed markers (gray circles) in stage 1 GWAS samples, and top SNPs in a combined analysis of all eight studies (red diamonds) are shown with $-\log_{10} P$ values (*Left y* axis). Overlaid are the likelihood-ratio statistics (*Right y* axis) to estimate putative recombination hotspots based on five sets of 100 randomly selected controls (connected lines in various colors). (*A*) Adjusted for study sites/regions, age, sex, and smoking status. (*B*) Additionally adjusted for rs2294008.

to 33 SNPs in the *PSCA* region genotyped in our stage 1 GWAS (14), we imputed 642 additional markers in the same 8,652 subjects. All markers passed the Hardy–Weinberg equilibrium test with a P value > 0.01 (Table S1). Of 675 markers in this region, 446 high-quality markers were used for further association analysis, including 413 SNPs with IMPUTE2-info score ≥ 0.9 and 33 SNPs already genotyped in the GWAS. Sequencing of all the *PSCA* exons in 45 bladder tumor samples did not find any coding variants not already included in our analyses.

The strongest signal was for rs2976393 [odds ratio (OR) = 1.11, 95% confidence interval (CI) = 1.04–1.18, $P = 1.62 \times 10^{-3}$] (Fig. 1*A* and Dataset S1). The original GWAS signal, rs2294008, ranked 22nd among these stage 1 samples with a per-allele OR of 1.09 (95% CI = 1.03–1.17, $P = 5.57 \times 10^{-3}$) (Fig. 1*A*). Among the 21 SNPs with a P value lower than that of rs2294008, 17 variants showed similar effect size and were in moderate-to-high linkage disequilibrium (LD) with rs2294008 (r^2 = 0.53–0.99) (Dataset S1). To test for the presence of other disease-associated SNPs in this region, we adjusted for the effect of rs2294008 along with age, sex, study regions, and smoking habit, and found the strongest signal for an additional variant, already genotyped in the GWAS, rs2978974 (OR = 1.11, 95% CI = 1.04–1.19, $P = 1.80 \times 10^{-3}$) (Fig. 1*B* and Dataset S1). Among 446 markers examined in this study, 46 SNPs are in strong LD ($r^2 > 0.8$) with rs2294008, but rs2978974 is a unique variant that is in low LD with rs2294008 ($D' = 0.19$, $r^2 = 0.02$) and not well tagged by other SNPs ($r^2_{max} = 0.19$) (Fig. S1 and Dataset S1). Additionally, rs2294008 and rs2978974 were genotyped in three independent studies, New England Bladder Cancer Study from New Hampshire (NEBCS-NH, a component of NEBCS not used in stage 1 GWAS), and Texas Bladder Cancer Study discovery set and validation set (TXBCS1 and TXBCS2, respectively). In the combined analysis of all eight studies that included 5,393 cases and 7,324 controls, rs2294008 and rs2978974 were associated with bladder cancer with a per-allele OR = 1.11 (95% CI = 1.06–1.17, $P = 5.79 \times 10^{-5}$) and OR = 1.07 (95% CI = 1.02–1.13, $P = 9.66 \times 10^{-3}$), respectively (Table 1).

Because rs2294008 and rs2978974 are in low LD but each showed evidence for association, we assessed the joint effect of these variants on bladder cancer risk (Table 2). Within the nine possible combinations of SNP genotypes, individuals carrying risk alleles contributed by both SNPs were at increased risk for bladder cancer, with individuals homozygous for risk alleles of both SNPs being at the highest risk (OR = 1.44, 95% CI = 1.08–1.94, $P = 1.48 \times 10^{-2}$) (Table 2). In a dominant model for the two SNPs, individuals carrying variants of both SNPs had increased risk of bladder cancer (OR = 1.24, 95% CI = 1.08–1.41, $P = 1.76 \times 10^{-3}$) (Table 2). There were nominally significant interactions between rs2294008 and rs2978974: multiplicative, P value of log-likelihood ratio test = 0.035 (Table 2) and additive, P value for relative excess risk for interaction = 0.039 (Table S2), also detected by a recently developed statistical method (22) (Table S2). The association results did not significantly differ by sex or tumor grade (Tables S3–S5). We observed an interaction between the two-SNP joint-effect model with smoking status (P = 0.027) (Table S5), with stronger bladder cancer risk among ever (vs. never) smokers carrying risk variants of both SNPs; however, this requires further confirmation studies.

***PSCA* mRNA Expression Analysis.** RNA-sequencing in normal and tumor bladder tissue samples detected 11 *PSCA* isoforms, but most of them were expressed at low levels (Fig. 2 and Table S6). *PSCA* was mainly found as a transcript that includes exons 1, 2, and 3 (*PSCA*-123, RefSeq NM_005672), with rs2294008 located within the first exon or 5′UTR of this transcript (Fig. 2). The other bladder cancer-associated SNP, rs2978974, is located 10 Kb upstream of rs2294008 within an alternative untranslated first exon (exon 1a) spliced to exon 2, creating an isoform *PSCA*-1a2 [GenBank NR_033343.1 (23)] (Fig. 2). This isoform is predicted to create a different ORF starting from exon 3 and generate a protein of 189 amino acids (Fig. S2), without similarity to PSCA or any other known protein.

Microarray expression analysis in 37 bladder tumors showed significant association for the risk T allele of rs2294008 with increased *PSCA* expression (p_{trend} = 0.0007) but there was no association for rs2978974 (p_{trend} = 0.1495) (Fig. 3*A*). We performed validation analysis with quantitative RT-PCR in 35 independent tumor bladder tissue samples using two custom-designed assays to detect the full-length *PSCA* transcript (*PSCA*-123) and a form with the alternative exon 1a (*PSCA*-1a2). *PSCA*-123 mRNA expression was strongly increased in individuals with risk T allele of rs2294008 (p_{trend} = 0.0054) but not of rs2978974 (p_{trend} = 0.9486) (Fig. 3*B*). A similar pattern of expression was observed in 35 normal bladder tissue samples (p_{trend} = 0.0155 for rs2294008 and p_{trend} = 0.7297 for rs2978974) (Fig. 3*C*). The genotype combination was not associated with *PSCA* expression (Fig. 3). Expression of the alternative form *PSCA*-1a2 was significantly lower than of *PSCA*-123, and it was associated with rs2294008 only in tumor samples (Fig. S3). There was significantly higher *PSCA* mRNA expression in tumors vs. normal bladder samples, with 7.3-fold (P = 0.0046) for

GENETICS

Table 1. Single marker analysis of rs2294008 and rs2978974 for association with bladder cancer risk

SNP	Cases n = 5,393 (n, %)	Controls n = 7,324 (n, %)	OR* (95% CI)	P value*	OR† (95% CI)	P value†	OR‡ (95% CI)	P value‡	OR§ (95% CI)	P value§
rs2294008										
CC	1,363 (25.27)	2,107 (28.77)	1.00 (–)	—	1.11 (1.06–1.17)	5.79E-05	1.00 (–)	—	1.13 (1.07–1.19)	1.07E-05
CT	2,804 (51.99)	3,645 (49.77)	1.19 (1.09–1.30)	7.81E-05			1.21 (1.11–1.32)	2.63E-05		
TT	1,226 (22.73)	1,572 (21.46)	1.23 (1.11–1.37)	1.07E-04			1.26 (1.13–1.40)	2.01E-05		
rs2978974										
GG	2,173 (40.29)	3,113 (42.50)	1.00 (–)	—	1.07 (1.02–1.13)	9.66E-03	1.00 (–)	—	1.09 (1.03–1.15)	1.63E-03
GA	2,475 (45.89)	3,325 (45.40)	1.05 (0.97–1.14)	2.12E-01			1.07 (0.99–1.15)	1.15E-01		
AA	745 (13.81)	886 (12.10)	1.17 (1.04–1.32)	7.55E-03			1.21 (1.08–1.36)	1.26E-03		

*Estimates from logistic regression models assuming genotypic effect, adjusted for study sites/regions, age, sex, and smoking status.
†Estimates from logistic regression models assuming log-additive genetic effect, adjusted for study sites/regions, age, sex, and smoking status.
‡Estimates from logistic regression models assuming genotypic effect, adjusted for study sites/regions, age, sex, and smoking status, and rs2294008/rs2978974.
§Estimates from logistic regression models assuming log-additive genetic effect, adjusted for study sites/regions, age, sex, smoking status, and rs2294008/rs2978974.

the *PSCA*-123 but not for the *PSCA*-1a2 form (Fig. 4). Analysis of 27 paired normal-tumor bladder tissue samples showed that the *PSCA*-123 expression was mainly affected by rs2294008 genotype but not the sample (normal-tumor) status (Fig. S4).

DNA–Protein Interaction. Although rs2978974 showed no association with *PSCA* mRNA expression, we hypothesized that it could have a regulatory effect because of its location within the first alternative untranslated exon. To further investigate the functional effect of this variant, we performed EMSA with allele-specific probes for rs2978974, using nuclear extracts from HeLa (cervix), LNCaP (prostate), LNCaP treated with dihydrotestosterone (DHT; a potent analog of testosterone), HTB-5 (bladder), and J82 (bladder) cancer cell lines. In all cell lines tested, there was a strong interaction with the probe for the non-risk G allele of rs2978974 (Fig. 5, lane 9), but not for the risk allele A (Fig. 5, lane 2). There were two distinct bands in LNCaP cells but only one band appeared in the other cell lines, suggesting a tissue-specific pattern for this DNA–protein interaction. It was also evident that the binding with the protein extract from DHT-treated LNCaP cells was much stronger than with protein extracts from all other cells, suggesting a specific role of testosterone in this regulation. Competition assays with excess of unlabeled rs2978974 G probe confirmed that both bands were specific (Fig. 5, lanes 10 and 11) because the intensities of these bands were decreased. However, the excess of unlabeled rs2978974 A probe did not cause the same effect (Fig. 5, lanes 12 and 13), confirming that the binding was specific for the G probe.

Two bioinformatic tools, AliBaba2 (24) and TFSearch (25), predicted interaction of the probe with the non-risk G allele of rs2978974 with proteins from the ETS family of transcription factors, specifically, with ELK1 (26), but no binding was predicted for the risk A allele (Fig. S5). After adding anti-ELK1 antibodies into binding reactions, we observed a moderate decrease of binding in the LNCaP cells, but there was no effect of ELK1 antibodies in HeLa, HTB-5, and J82 cell lines (Fig. 5, lane 14).

Discussion

The GWAS association between SNP rs2294008 within the *PSCA* gene and bladder cancer susceptibility was originally reported by Wu et al. (15) and has been subsequently confirmed by Rothman et al. (14), in which rs2294008 had a per-allele OR = 1.13 (95% CI = 1.09–1.17, $P = 4.4 \times 10^{-11}$) in the combined set of 10,196 cases and 44,705 controls. Here, we used imputation and selected genotyping followed by a conditional association analysis to fine-map this region. We confirmed that rs2294008 captured the main association with bladder cancer susceptibility in the *PSCA* region. Based on fine-mapping, we provide evidence for a unique signal detected by rs2978974, located 10 Kb upstream of rs2294008. We found that there was statistically significant interaction between the two SNPs, with an increased bladder cancer risk among carriers of risk alleles from both SNPs. This finding does, however, require replication in future studies. Higher *PSCA* mRNA expression was detected in bladder tumor samples compared with adjacent normal bladder tissue, with enhanced expression specifically present among carriers of the risk allele T of rs2294008. Although no effect on mRNA expression was observed for rs2978974, we detected

Table 2. Joint effects analysis of rs2294008 and rs2978974 for association with bladder cancer risk

rs2294008	rs2978974	Cases n = 5,393(n, %)	Controls n = 7,324 (n, %)	OR* (95% CI)	P value*	OR_{inter} (95% CI)	P $value_{inter}$§
CC	GG	477 (8.84)	716 (9.78)	1.00 (–)	—	1.05 (0.97–1.14)†	0.1967†
CC	AG	620 (11.50)	1,035 (14.13)	0.91 (0.77–1.06)	2.19E-01		
CC	AA	266 (4.93)	356 (4.86)	1.08 (0.88–1.33)	4.41E-01		
CT	GG	1,094 (20.29)	1,579 (21.56)	1.05 (0.91–1.21)	5.40E-01		
CT	AG	1,342 (24.86)	1,653 (22.57)	1.21 (1.05–1.40)	7.50E-03		
CT	AA	368 (6.82)	413 (5.64)	1.33 (1.10–1.60)	3.36E-03		
TT	GG	602 (11.16)	818 (11.17)	1.15 (0.98–1.35)	9.93E-02		
TT	AG	513 (9.51)	637 (8.70)	1.20 (1.01–1.42)	3.52E-02		
TT	AA	111 (2.06)	117 (1.60)	1.44 (1.08–1.94)	1.48E-02		
CC	GG	477 (8.84)	716 (9.78)	1.00 (–)	—	1.20 (1.01–1.42)‡	0.0354‡
CC	AG+AA	886 (16.43)	1,391 (18.99)	0.95 (0.82–1.10)	5.16E-01		
CT+TT	GG	1,696 (31.45)	2,397 (32.73)	1.08 (0.94–1.24)	2.68E-01		
CT+TT	AG+AA	2,334 (43.28)	2,820 (38.50)	1.24 (1.08–1.41)	1.76E-03		

*Estimates from logistic regression models, adjusted for study sites/regions, age, sex, and smoking status, based on the same reference group.
†Multiplicative interaction estimates assuming log-additive effects for both SNPs, adjusted for study sites/regions, age, sex, and smoking status.
‡Multiplicative interaction estimates assuming dominant effects for both SNPs, adjusted for study sites/regions, age, sex, and smoking status.
§Multiplicative interaction estimates (OR, 95% CI) assuming genotypic effects for both SNPs, adjusted for study sites/regions, age, sex, and smoking status are as follows: $Inter_{CT/AG}$ = 1.28 (1.06–1.55); $Inter_{CT/AA}$ = 1.17 (0.90–1.52); $Inter_{TT/AG}$ = 1.16 (0.92–1.45); $Inter_{TT/AA}$ = 1.16 (0.82–1.66).

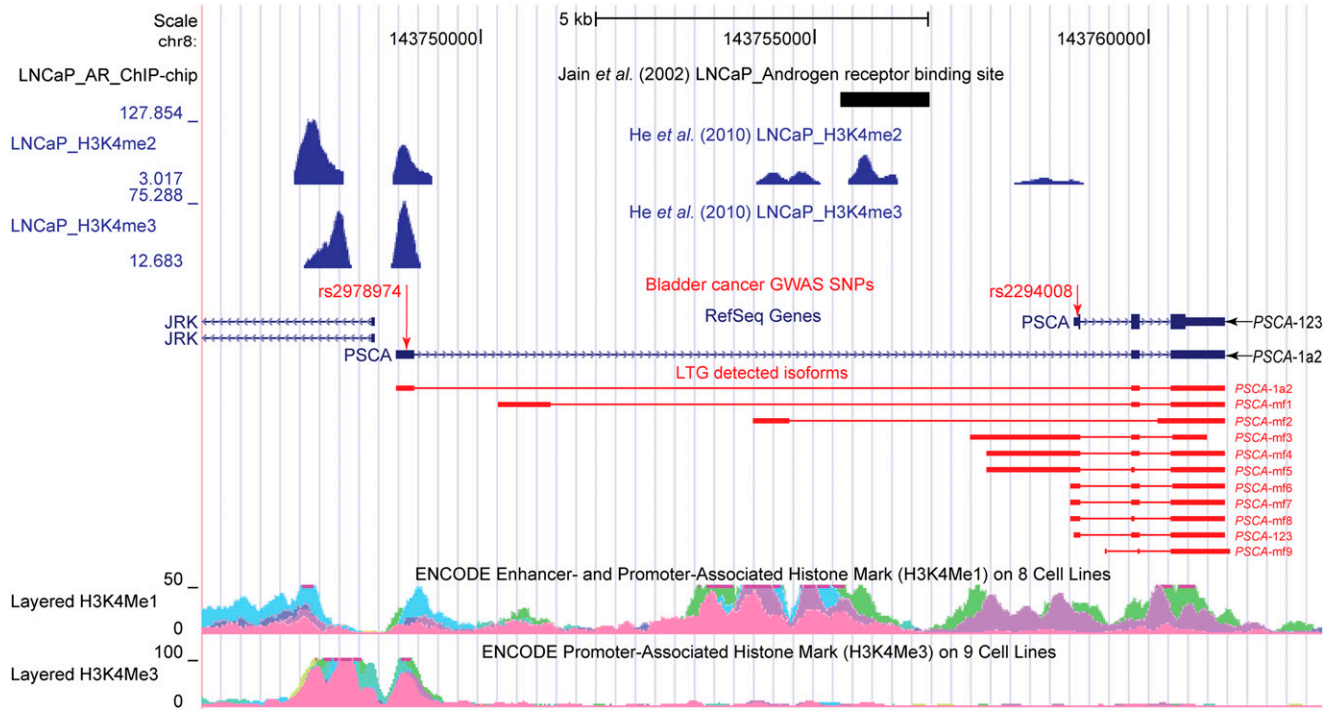

Fig. 2. RNA-sequencing results of seven bladder tumors and seven adjacent normal bladder tissue samples in the *PSCA* region. Data for androgen receptor binding sites (37) and histone methylation marks (27, 28) was obtained from publicly available datasets.

a strong and specific binding of nuclear proteins from five human cancer cell lines to the non-risk allele G. Absence of similar binding to the risk A allele of rs2978974 suggests the disruption of a regulatory function as a possible mechanism for this risk variant. Although it seems less likely that these two SNPs affect the same molecular mechanism, such as mRNA expression, they might be involved in two independent mechanisms that contribute to increased bladder cancer risk.

Both rs2294008 and rs2978974 have differences in minor allele frequencies in distinct populations (Table S7), reflected in population-specific LD patterns in this region (Fig. S1). All of the subjects in our GWAS were evaluated for differences in underlying population substructure using population-specific genetic markers, and samples with significant non-European admixture were excluded; additional follow-up samples were of self-described ethnicity. The observed heterogeneity (Fig. S6) might be the result of higher population admixture in some of the studies or because of chance alone.

Although the risk allele T of rs2294008 was associated with increased mRNA expression both in bladder tumor and adjacent normal tissue samples, this functional effect might be attributed to any of the 46 variants that were in high LD ($r^2 > 0.8$) with rs2294008 (Fig. 1 and Dataset S1). The risk allele T of rs2294008 creates a novel translation start site nine amino acids upstream of the regular start site, extending PSCA leader peptide from 11 to 20 amino acids (Fig. S2). The functional significance of this extension of the PSCA protein is currently under active evaluation.

In contrast with rs2294008, rs2978947 is a distinct variant not in strong LD with other markers ($r^2_{max} = 0.19$; Dataset S1). We located rs2978974 in an untranslated alternative first exon of the *PSCA* gene, 10 Kb upstream of rs2294008. This alternative exon is spliced directly to exon 2 of *PSCA*, generating a low-expressing mRNA transcript that is predicted to encode a protein unrelated to PSCA. In the published datasets (27, 28), the area surrounding rs2978974 contains strong epigenetic marks represented by methylation of histone 3 at lysine 4 (H3K4me2 and H3K4me3), which are often found to be associated with active regulatory elements such as promoters and enhancers (29). Because rs2978974 is a noncoding SNP located in an alternative untranslated first exon of the *PSCA* gene, it is reasonable to speculate that this SNP might function as a regulatory element for *PSCA* through allele-specific binding of transcription factors (30, 31).

We predicted that the non-risk G allele of rs2978974, but not the risk allele A, could bind ETS transcription factors, and ELK1, specifically. The 27 known ETS proteins modulate cell proliferation, differentiation, migration, invasiveness, and response to the microenvironment, which are the factors important for growth and metastasis of solid tumors (32, 33). ELK1 is specifically involved in the transcriptional regulation by forming a ternary complex with serum response factors at the *c-fos* serum response element (34), and has been shown as a regulator for vascular smooth muscle differentiation in the murine bladder (35). Our EMSA results with nuclear extracts from five cell lines confirmed the strong and allele-specific interaction with nuclear proteins, but the supershift assays with an anti-ELK1 antibody showed only a weak effect in the LNCaP prostate cancer cell line. It is possible that other ETS proteins are also involved in this regulation, and they are differentially expressed in prostate compared with bladder cancer cells. Recently, ELK1 expression was found elevated in an androgen-sensitive cell line LNCaP, but not in androgen-insensitive cell lines PC3 and DU-145 (36). We also observed that the allele-specific binding of rs2978974 was stronger in the DHT-treated cell line LNCaP compared with nontreated cells (Fig. 5). There is an androgen receptor binding site located between rs2294008 and rs2978974 within the *PSCA* promoter region (Fig. 2) (37), but the exact role of testosterone in regulation of *PSCA* function is still unclear.

In summary, using the results of a bladder cancer GWAS, we have identified SNP rs2978974 as an additional marker for bladder cancer susceptibility in the *PSCA* region of 8q24.3. The joint effect of risk alleles from the original GWAS signal rs2294008 and this recently identified marker rs2978974 suggests that both variants contribute to bladder cancer susceptibility. The risk allele A of rs2978974 might be related to the loss of binding to ELK1 or other ETS proteins, and the functional significance of rs2294008 or another variant in high LD with it could contribute to differential regulation of *PSCA* mRNA expression.

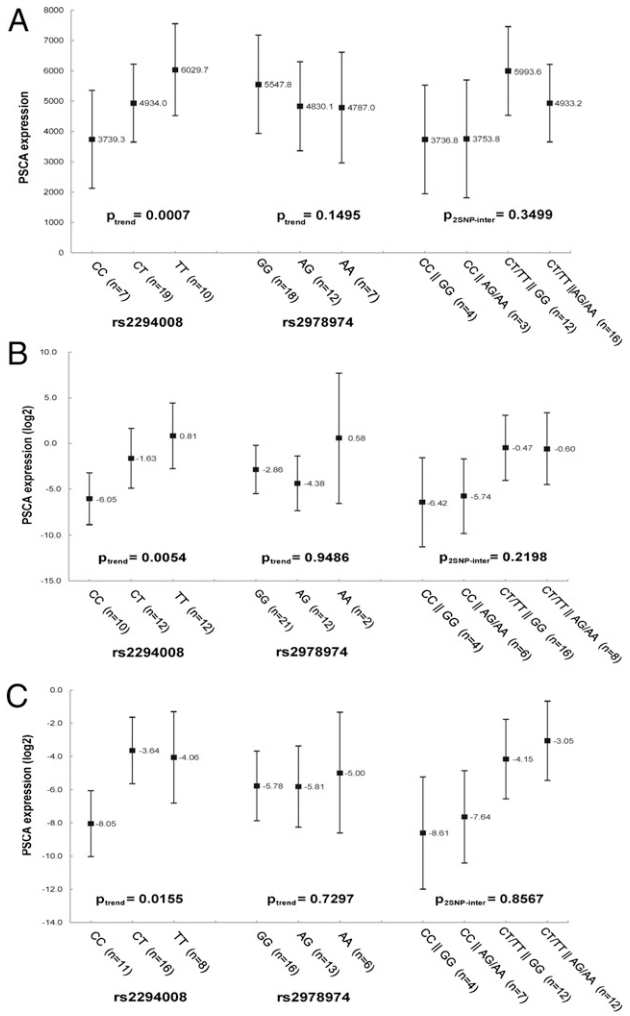

Fig. 3. *PSCA* mRNA expression stratified by rs2294008 and rs2978974 genotypes. Expression is shown with 95% CI, estimated from generalized linear models adjusted for the effect of age and sex. Total *PSCA* expression (*A*) in bladder tumor tissues measured with microarrays, (*B*) in bladder-tumor tissues measured with an expression assay *PSCA*-123, and (*C*) in normal-bladder tissues measured with an expression assay *PSCA*-123.

Materials and Methods

Study subjects, genotyping, imputation, and association analyses are described in *SI Materials and Methods*. IMPUTE2 (38) was used to impute additional SNP genotypes based on the combined reference panel from the 1000 Genomes Project [June 2010 release (39)] and HapMap Phase 3 CEU data [February 2009 release 2 (40)]. The estimated allelic dosage for each imputed marker was used in SNPTEST v2 (41) to test for association with bladder cancer risk.

Cell Lines and Tissue Samples. Cell lines were from American Type Culture Collection. Fresh-frozen bladder tissue samples were obtained from the Spanish Bladder Cancer Case-Control Study EPICURO (42), collected after approval by the National Cancer Institute Institutional Review Board and the ethics committees of all participating hospitals or from Asterand, purchased after exemption #4715 by the National Institutes of Health Office of Human Subjects Research. All samples were processed as described in *SI Materials and Methods*.

mRNA Expression Analysis. RNA-sequencing (RNA-seq) of bladder tissue samples was performed by the National Cancer Institute Core Sequencing Facility. *PSCA* mRNA expression was analyzed with the Human Gene ST1.0 expression array (Affymetrix) for the first set and with TaqMan expression assays for the second set of bladder tissue samples, as described in *SI Materials and Methods*.

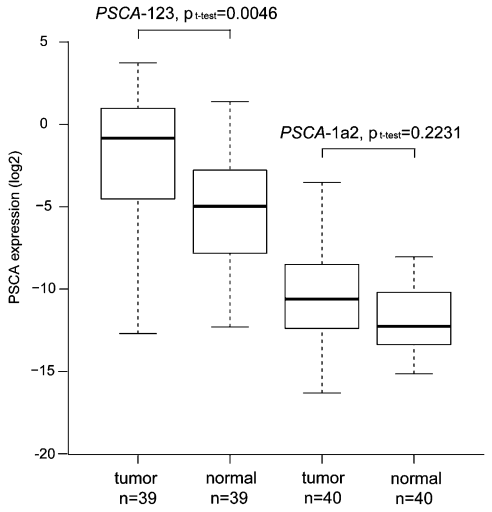

Fig. 4. mRNA expression of two *PSCA* isoforms in tumor and normal bladder tissues.

DNA–Protein Interaction Analysis. Bioinformatic prediction analyses were performed with AliBaba2 (24) and TFSearch (25). Experimental testing of DNA-protein interactions for rs2978974 was performed with EMSA, as described in *SI Materials and Methods*.

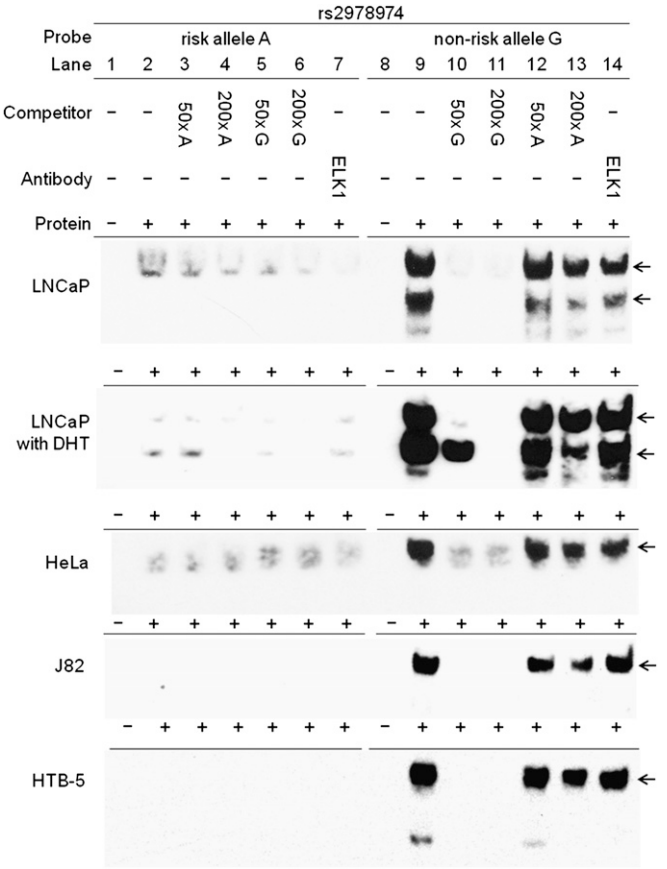

Fig. 5. EMSA results for SNP rs2978974 in five cancer cell lines: LNCaP, LNCaP treated with testosterone analog DHT, HeLa, J82, and HTB-5. The shifted bands (lane 9) indicating DNA-protein interactions were only detected for non-risk G allele (*Right*) but not for the risk allele A (*Left*), and this allele-specific binding was supported by competition assays (lane 12 and 13). Effects on binding of an anti-ELK1 antibody were shown in lane 14.

ACKNOWLEDGMENTS. We thank the following for their support of this project: Leslie Carroll, Jane Wang, Anne Taylor, and Kirk Snyder (Information Management Services); Gemma Castaño-Vinyals, Fernando Fernández, Maria Sala, and Montserrat Torà (Institut Municipal d'Investigació Mèdica); Paul Hurwitz, Charles Lawrence, Anna McIntosh, and Robert Saal (Westat, Inc.); Marta Lopez-Brea and Fernando Rivera (Marqués de Valdecilla University Hospital); and Angeles Panadero (Hospital Ciudad de Coria). This project has been funded in part by the Intramural Research Program of the National Cancer Institute, National Institutes of Health, under Contract HHSN261200800001E. Additionally, this research was supported by US Public Health Service Contracts N01-CN-45165, N01-RC-45035, and N01-RC-37004 from the National Cancer Institute, Department of Health and Human Services. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Additional funding information is provided in *SI Materials and Methods*.

1. Jemal A, et al. (2011) Global cancer statistics. *CA Cancer J Clin* 61:69–90.
2. Siegel R, Naishadham D, Jemal A (2012) Cancer statistics, 2012. *CA Cancer J Clin* 62: 10–29.
3. American Cancer Society (2012) Cancer Facts and Figures 2012. Available at http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-031941.pdf. Accessed January 2012.
4. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R (2003) The health economics of bladder cancer: A comprehensive review of the published literature. *Pharmacoeconomics* 21:1315–1330.
5. Avritscher EB, et al. (2006) Clinical model of lifetime cost of treating bladder cancer and associated complications. *Urology* 68:549–553.
6. Silverman DT, Devesa SS, Moore LE, Rothman N (2006) Bladder Cancer. *Cancer Epidemiology and Prevention,* Chapter 58, eds Schottenfeld D, Fraumeni JF (Oxford Univ Press, New York), pp 1101–1127.
7. Lichtenstein P, et al. (2000) Environmental and heritable factors in the causation of cancer—Analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343:78–85.
8. Czene K, Lichtenstein P, Hemminki K (2002) Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. *Int J Cancer* 99:260–266.
9. Stern MC, et al.; International Consortium of Bladder Cancer (2009) Polymorphisms in DNA repair genes, smoking, and bladder cancer risk: Findings from the international consortium of bladder cancer. *Cancer Res* 69:6857–6864.
10. García-Closas M, et al. (2005) NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: Results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 366:649–659.
11. Rothman N, Garcia-Closas M, Hein DW (2007) Commentary: Reflections on G. M. Lower and colleagues' 1979 study associating slow acetylator phenotype with urinary bladder cancer: meta-analysis, historical refinements of the hypothesis, and lessons learned. *Int J Epidemiol* 36:23–28.
12. Figueroa JD, Garcia-Closas M, Rothman N (2010) Chapter 15. Case studies: Cumulative assessment of the role of human genome variation in specific diseases - bladder cancer. *Human Genome Epidemiology (2nd Edition): Building the Evidence for Using Genetic Information to Improve Health and Prevent Disease*, eds Khoury MJ, et al. (Oxford Univ Press, New York), pp 299–325.
13. Kiemeney LA, et al. (2008) Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. *Nat Genet* 40:1307–1312.
14. Rothman N, et al. (2010) A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. *Nat Genet* 42:978–984.
15. Wu X, et al. (2009) Genetic variation in the prostate stem cell antigen gene *PSCA* confers susceptibility to urinary bladder cancer. *Nat Genet* 41:991–995.
16. Garcia-Closas M, et al. (2011) 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* 20:4282–4289.
17. Tang W, et al. (2012) Mapping of the UGT1A locus identifies an uncommon coding variant that affects mRNA expression and protects from bladder cancer. *Hum Mol Genet*, 10.1093/hmg/ddr619.
18. Reiter RE, et al. (1998) Prostate stem cell antigen: A cell surface marker overexpressed in prostate cancer. *Proc Natl Acad Sci USA* 95:1735–1740.
19. Antonarakis ES, Carducci MA, Eisenberger MA, Denmeade SR, Slovin SF, et al. (2012) Phase I rapid dose-escalation study of AGS-1C4D4, a human anti-PSCA (prostate stem cell antigen) monoclonal antibody, in patients with castration-resistant prostate cancer: A PCCTC trial. *Cancer Chemother Pharmacol* 69:763–771.
20. Wente MN, et al. (2005) Prostate stem cell antigen is a putative target for immunotherapy in pancreatic cancer. *Pancreas* 31:119–125.
21. Chung CC, Chanock SJ (2011) Current status of genome-wide association studies in cancer. *Hum Genet* 130:59–78.
22. Han SS, Rosenberg PS, Garcia-Closas M, Figueroa JD, Silverman D, et al. (2012) Likelihood ratio test for detecting gene (G) environment (E) interactions under the additive risk model exploiting G-E independence for case-control data. *Am J Epidemiol*, in press.
23. Geiger KD, et al. (2011) The prostate stem cell antigen represents a novel glioma-associated antigen. *Oncol Rep* 26:13–21.
24. Grabe N (2002) AliBaba2: Context specific identification of transcription factor binding sites. *In Silico Biol* 2:S1–S15.
25. TFSEARCH Searching Transcription Factor Binding Sites. Available at http://www.cbrc.jp/research/db/TFSEARCH.html. Accessed December 2011.
26. Yamauchi T, Toko M, Suga M, Hatakeyama T, Isobe M (1999) Structural organization of the human Elk1 gene and its processed pseudogene Elk2. *DNA Res* 6:21–27.
27. He HH, et al. (2010) Nucleosome dynamics define transcriptional enhancers. *Nat Genet* 42:343–347.
28. Raney BJ, et al. (2011) ENCODE whole-genome data in the UCSC genome browser (2011 update). *Nucleic Acids Res* 39(Database issue):D871–D875.
29. Santos-Rosa H, et al. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* 419:407–411.
30. Birney E, et al.; ENCODE Project Consortium; NISC Comparative Sequencing Program; Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447:799–816.
31. Visel A, Rubin EM, Pennacchio LA (2009) Genomic views of distant-acting enhancers. *Nature* 461:199–205.
32. Hsu T, Trojanowska M, Watson DK (2004) Ets proteins in biological control and cancer. *J Cell Biochem* 91:896–903.
33. Charlot C, Dubois-Pot H, Serchov T, Tourrette Y, Wasylyk B (2010) A review of post-translational modifications and subcellular localization of Ets transcription factors: Possible connection with cancer and involvement in the hypoxic response. *Methods Mol Biol* 647:3–30.
34. Hipskind RA, Rao VN, Mueller CG, Reddy ES, Nordheim A (1991) Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62TCF. *Nature* 354:531–534.
35. Li J, et al. (2006) Serum response factor, its cofactors, and epithelial-mesenchymal signaling in urinary bladder smooth muscle formation. *Differentiation* 74:30–39.
36. Shaikhibrahim Z, Lindstrot A, Langer B, Buettner R, Wernert N (2011) Differential expression of ETS family members in prostate cancer tissues and androgen-sensitive and insensitive prostate cancer cell lines. *Int J Mol Med* 28:89–93.
37. Jain A, Lam A, Vivanco I, Carey MF, Reiter RE (2002) Identification of an androgen-dependent enhancer within the prostate stem cell antigen gene. *Mol Endocrinol* 16: 2323–2337.
38. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 5: e1000529.
39. Durbin RM, et al.; 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467:1061–1073.
40. The International HapMap3 Consortium (2010) Integrating common and rare genetic variation in diverse human populations. *Nature* 467:52–58.
41. SNPTEST v2. Available at https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html. Accessed December 2011.
42. Samanic C, et al. (2006) Smoking and bladder cancer in Spain: Effects of tobacco type, timing, environmental tobacco smoke, and gender. *Cancer Epidemiol Biomarkers Prev* 15:1348–1354.