Supplementary Methods

Tissue samples and cell lines

The previously described (1) fresh-frozen tumor (n=42) and adjacent normal (n=39) bladder tissues samples and 10 normal-tumor paired formalin fixed paraffin embedded (FFPE) blocks were purchased from Asterand (Detroit, MI) after exemption #4715 by the NIH Office of Human Subjects Research (OHSR). Additional samples of bladder tumors were from The New England Bladder Cancer Study (NEBCS) (1) further described in immunohistochemistry section. HeLa cell line was purchased from the American Type Culture Collection (ATCC, Manassas, VA) and maintained according to recommended conditions.

Allelic Expression Imbalance (AEI)

RNA-sequencing of 6 tumor and 6 adjacent normal bladder tissue samples has been previously described (2). RNA-seq reads corresponding to *PSCA* region were extracted from the alignment files and visualized by Integrative Genomics Viewer (http://www.broadinstitute.org/igv.). AEI was revealed by counting of RNA-seq reads carrying T and C alleles of rs2294008. Location of rs2294008 within the first exon of *PSCA* allowed the use of the same rs2294008 allelic discrimination genotyping assay (C_3288106_1, Applied Biosystems) both for DNA and cDNA samples. AEI was tested in normal/tumor bladder tissue samples heterozygous for rs2294008 as determined by DNA genotyping. For each sample, cDNA was synthesized from 1µg of total RNA using SuperScript III kit (Invitrogen) and random hexamers. AEI experiments were performed using post-read analysis of the cDNA expression plates. The ratio between expression in tumor and normal tissue (T:N) was calculated for both alleles in DNA and cDNA samples as the post-run allelic discrimination ratio of VIC (allele C)

and FAM (allele T) fluorophores, labeling corresponding alleles of rs2294008. The assay was initially validated by creating a standard curve representing serial dilution of two DNA samples with TT and CC genotype. Based on 8 dilutions at the specific ratios (5, 15, 25, 35, 45, 55, 65, 75, 85% of one of the alleles), we constructed a standard curve with r=0.94, confirming high performance and sensitivity of this assay to quantify both alleles of rs2294008 (Supplementary Table 2).

cDNA cloning

Full-length *PSCA* transcripts carrying the C and T alleles of rs2294008 were amplified from bladder cDNA and cloned into the pFC14A (Halo-tag) CMV Flexi vector (Promega) to create fusion proteins with C-terminal Halo-tags. A common reverse primer- PmeI_R: 5'-gattgtttaaacGAGCTGGCCGGGTCCCCAGAGCAGC-3' and two allele-specific forward primers: *PSCA*_C_SgfI_F: 5'- aatcgcgatcgcc**ATG**GCAGGCTTGGCCCTGCAGCCAG-3' and *PSCA*_T_SgfI_F: 5'- aatcgcgatcgcc**ATG**AAGGCTGTGCTGCTTGCCCT-3' were used for cloning. The plasmids were verified by automated sequencing on 3730 Sequencer (Applied Biosystems) and alignment with Sequencher 4.8 software (Gene Code).

Antibody

Proteolytic removal of the N-terminal leader peptide and the C-terminal part of protein during attachment of the GPI anchor complicates the detection of PSCA expression with antibodies for the N or C-terminal fusion proteins. In this work we used a mouse monoclonal anti-PSCA antibody (1G8) (3, 4) kindly provided by Dr. Robert Reiter (UCLA). The antibody was raised against amino acids 46-85 of the human PSCA protein, the area not affected by any genetic variation. Thus, the antibody should be detecting both allelic PSCA protein isoforms. The antibody was previously demonstrated to preferentially detect PSCA expression on the cell

surface, however, denatured PSCA protein was not well detectable by this antibody by Western blotting (5, 6). We found this antibody to provide the best imaging results in our pilot IHC experiment, compared to several other antibodies tested - mouse monoclonal PSCA 7F5, sc-80654 (Santa Cruz) and H00008000-B01P (Abnova); and rabbit polyclonal antibody ab64919 (Abcam) (data not shown). PSCA protein expression using 1G8 antibody was confirmed in HeLa cell line transiently transfected with both PSCA allelic expression constructs. A humanized version of this antibody (AGS-1C4D3, Agensys, Astellas) is used for anti-PSCA immunotherapy.

Fluorescence activated cell sorting (FACS)

HeLa cells were transfected with PSCA allelic expression constructs and control constructs in 6-well plates, in triplicates, using Lipofectamine™ LTX Reagent (Invitrogen). Non-transfected cells (Lipofectamine only) and cells transfected with empty Halo-tag control vector (Promega) were used as negative controls. Cell media was changed 24 hours post-transfection. After 48 hours, cells were collected using enzyme free cell dissociation buffer (Invitrogen) and incubated with 1G8 anti-PSCA antibody (0.39 µg/mL) for 45 minutes on ice, followed by incubation with anti-mouse Cy5 labeled secondary antibody (ab6785) for 30 minutes at a 1:1500 dilution. Samples were resuspended in 300µl of cell staining buffer (BioLegends) containing DAPI at 1:5000 dilution and incubated on ice in dark for 5 minutes. FACS analysis was performed on FACS AriaII (BD Biosciences). FACS analysis with isotype control (IgGk1; ab18443, Abcam) was used as a negative control. Transfection with an empty GFP vector was used to monitor the transfection efficiency. Data was analyzed using FlowJo software (Tree Star, Inc., Ashland, OR). Student's T-test was used to compare the percentage of positive cells between samples transfected with PSCA-T construct and those transfected with

PSCA-C construct, as well as between cells transfected with PSCA-C or PSCA-T construct and mock controls.

Immunohistochemistry (IHC) analysis

A pilot custom tissue microarray (TMA) was generated by Asterand Inc. (Detroit). The TMA included 20 bladder tissue samples - 5 normal/tumor pairs with CC genotype and 5 pairs with TT genotype of rs2294008. The samples were selected solely based on rs2294008 genotype determined in corresponding DNA samples. These bladder tissue samples were from patients with muscle-invasive bladder cancer (stage T2-T3). As positive controls for PSCA expression we included 3 high-grade prostate tumor samples (Gleason score >7). An independent set included 278 incident bladder cancer cases from Main and Vermont (The New England Bladder Cancer Study, NEBCS) (1, 7). The study samples and the construction of tissue microarrays (TMAs) were previously described (7). Briefly, these samples were from patients who provided written informed consent for collection of pathology materials and had available rs2294008 genotype data in GWAS (8). The study was approved by the Intramural Review Board of NIH and corresponding study centers. Formalin-fixed paraffin-embedded blocks of bladder tumor tissues were collected from the pathology archives together with tumor histopathologic characteristics (diagnosis as urothelial carcinoma, tumor grade and stage), independently confirmed by the study pathologist (A. Schned).

The TMAs were constructed using a Beecher MTA-1 (Beecher Instruments, Silver Spring, MD). Each sample was represented by a single 1-mm tissue core; 5μm tissue sections were cut, placed on glass slides using a tape transfer method, and UV-cross linked (Instrumedics, Hackensack, NJ). Slides were de-paraffinized, treated with “Superblock (Scytek) for 5 minutes with antigen retrieval by steaming for 5 minutes in a Dako Target Retrieval Solution (pH 9).

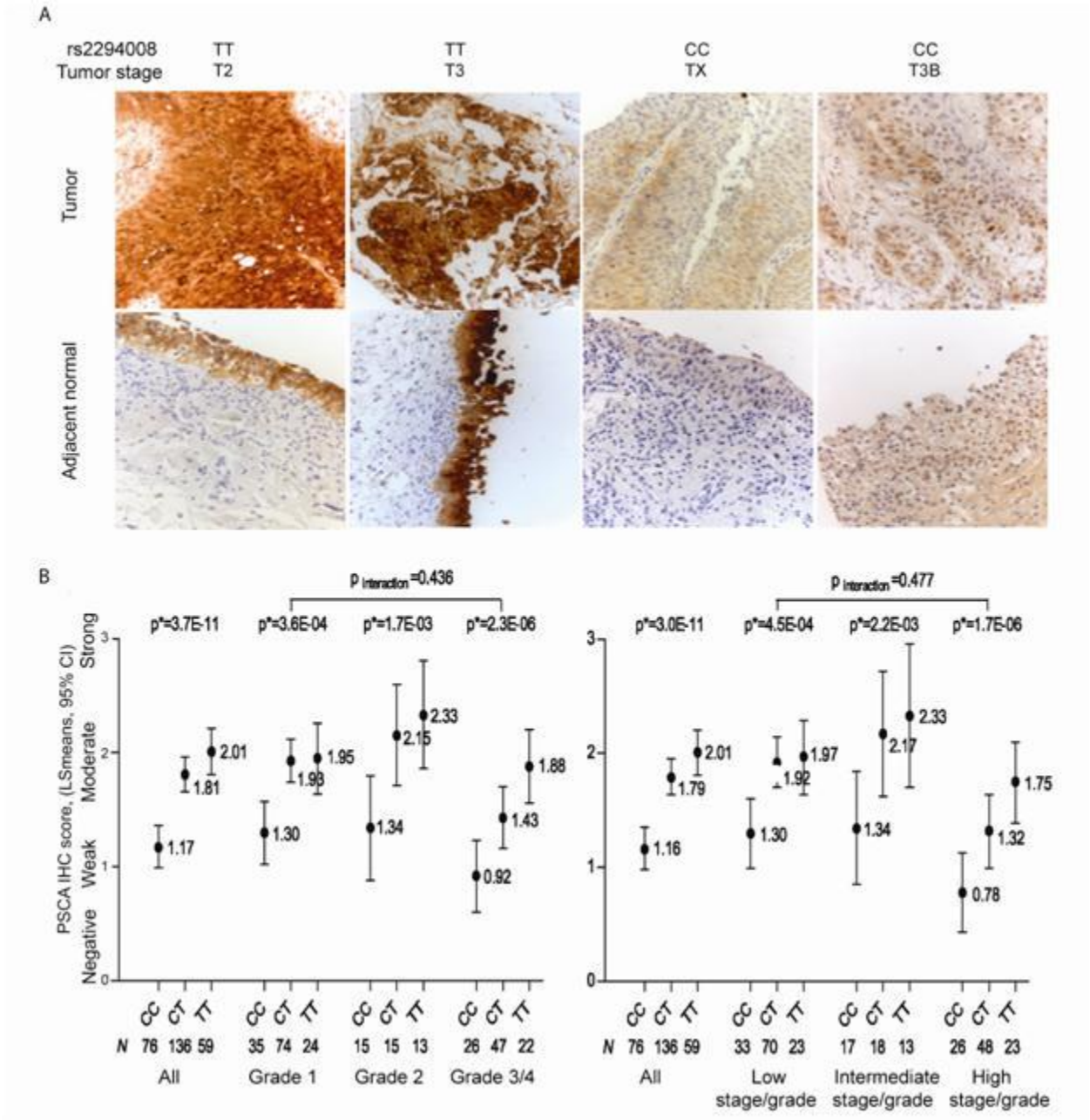
Primary monoclonal anti-PSCA antibody (clone 1G8) was applied at 1:630 dilution for 30 minutes and the reaction was developed with Envision+ (Dako). Slides were counterstained with hematoxylin, dehydrated, and coverslipped. Digital images were created by scanning whole slides using Aperio CS instrument with a 20X objective (Aperio Technologies, Vista, CA). PSCA expression was manually scored as negative (score 0), weak (score 1), moderate (score 2) and strong expression (score 3). Scoring was performed by a trained pathologist blindly to the genotype and clinical variables.

Statistical analysis

Generalized linear model was used to examine the association between PSCA IHC scores as continuous variables (0, 1, 2, 3) and rs2294008 genotypes. We assumed SNP rs2294008 follows an additive genetic model and coded the genotypes CC/CT/TT as 0, 1 and 2 based on the counts of the risk T allele. We also performed ordered logistic regression where PSCA IHC score was treated as an ordinal variable. Both models were adjusted for the effect of age, gender, study sites (Vermont/Maine) smoking habits (never/ever smokers), and tumor grade/stage. Stratified analyses and test for heterogeneity were performed for tumor stage, grade and stage/grade combined as a variable. The combined tumor stage/grade variable was defined into three groups: low stage/grade comprised all Ta cases with histology grade I; intermediate stage/grade comprised Ta cases with histology grade II, or T1/T2 cases with histology grades I/II; and high stage/grade comprised all cases with either stages T3/T4 or histology grade III. All statistical tests were two-sided and conducted with SAS/STAT system version 9.2 (SAS Institute Inc.), the graphs were plotted with Prism 5 software (GraphPad, La Jolla, CA).

References

1. Baris D, Karagas MR, Verrill C, Johnson A, Andrew AS, Marsit CJ, et al. A case-control study of smoking and bladder cancer risk: emergent patterns over time. *J Natl Cancer Inst* 2009;101(22):1553-61.
2. Fu YP, Kohaar I, Rothman N, Earl J, Figueroa JD, Ye Y, et al. Common genetic variants in the PSCA gene influence gene expression and bladder cancer risk. *Proc Natl Acad Sci U S A* 2012;109(13):4974-9.
3. Gu Z, Yamashiro J, Kono E, Reiter RE. Anti-prostate stem cell antigen monoclonal antibody 1G8 induces cell death in vitro and inhibits tumor growth in vivo via a Fc-independent mechanism. *Cancer Res* 2005;65(20):9495-500.
4. Zhennan Gu TO, Mark Sherman, Sherie L. Morrison, Anna M. Wu, Robert E. Reiter. Biological activity and microPET imaging properties of chimeric and humanized anti-prostate stem cell antigen (PSCA) antibodies. *Proceedings of the 96th Annual Meeting of the American Association for Cancer Research* 2005.
5. Gu Z, Thomas G, Yamashiro J, Shintaku IP, Dorey F, Raitano A, et al. Prostate stem cell antigen (PSCA) expression increases with high gleason score, advanced stage and bone metastasis in prostate cancer. *Oncogene* 2000;19(10):1288-96.
6. Saffran DC, Raitano AB, Hubert RS, Witte ON, Reiter RE, Jakobovits A. Anti-PSCA mAbs inhibit tumor growth and metastasis formation and prolong the survival of mice bearing human prostate cancer xenografts. *Proc Natl Acad Sci U S A* 2001;98(5):2658-63.
7. Lenz P, Pfeiffer RM, Baris D, Schned A, Takikita M, Poscablo MC, et al. Cell Cycle Control in Urothelial Carcinoma: Large-scale Tissue Array Analysis of Tumor Tissue from Maine and Vermont. *Cancer Epidemiol Biomarkers Prev*, 2012.
8. Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. *Nat Genet* 2010; 42(11):978-84.


A
rs2294008
Tumor stage
TT
T2
TT
T3
CC
TX
CC
T3B
Tumor
Adjacent normal
B
p interaction=0.436
p*=3.7E-11
p*=3.6E-04
p*=1.7E-03
p*=2.3E-06
PSCA IHC score, (LSmeans, 95% CI)
Strong
Moderate
Weak
Negative
1.17
1.81
2.01
1.30
1.93
1.95
1.34
2.15
2.33
0.92
1.43
1.88
CC CT TT CC CT TT CC CT TT CC CT TT
N 76 136 59 35 74 24 15 15 13 26 47 22
All
Grade 1
Grade 2
Grade 3/4
p interaction=0.477
p*=3.0E-11
p*=4.5E-04
p*=2.2E-03
p*=1.7E-06
1.16
1.79
2.01
1.30
1.92
1.97
1.34
2.17
2.33
0.78
1.32
1.75
CC CT TT CC CT TT CC CT TT CC CT TT
N 76 136 59 33 70 23 17 18 13 26 48 23
All
Low stage/grade
Intermediate stage/grade
High stage/grade

Supplementary Figure 1. Analysis of PSCA IHC expression by bladder tissue microarrays (TMAs).

A. Analysis of PSCA expression by IHC on bladder cancer TMA with 1G8 anti-PSCA antibody in 4 representative pairs of tumor-adjacent normal bladder tissue samples with TT and CC genotypes of rs2294008. In carriers of the risk TT genotype PSCA expression was detected even in normal tissue, while no or low expression was detected in carriers of the non-risk CC genotype even in tumor tissue.

B. Association of PSCA IHC scores and rs2294008 genotypes in 271 bladder tumors, in relation to tumor subgroups. *p-values are estimated from generalized linear models, assuming additive genetic effect of rs2294008, adjusted for age, sex, study site and smoking status (ever/never). Shown are mean values of corresponding groups with 95% CI.

Supplementary Table 1. Allelic Expression imbalance (AEI) for SNPs in *JRK, PSCA* and *LY6K* gene in normal-tumor bladder tissue samples

Gene	rs#	Location, hg19	SNP	%*											
				T_1	N_1	T_2	N_2	T_3	N_3	T_4	T_5	N_5	T_6	N_6	T_7	%†
JRK	rs1137522	143739020	**A:C**					51	46	39					45	45.3
	rs754957	143746416	**A:G**	75	78			52	44						64	62.6
	rs3802232	143746701	**A:G**	78	30			44	20	50					67	48.2
	rs2978973	143747271	**C:G**					50	50						52	50.7
PSCA	rs2294008	143761931	**T:C**					88	88	88				95	92	90.2
	rs2978982	143763490	**C:T**					92	90	94				93	93	92.4
	rs1045531	143763547	**A:C**					91	91	94			99	95	92	93.7
	rs2976393	143763618	**G:C**					93	92	96			99	95	95	95.0
	rs2976394	143763622	**T:C**					93	91	94			99	95	93	94.2
	rs10216533	143763690	**A:G**					91	89	91			99	92	91	92.2
	rs2976395	143763750	**A:G**					93	92	94				97	94	94.0
	rs1045547	143763757	**G:T**					94	93	97				98	95	95.4
	rs1045574	143763958	**A:G**					95	94	96				97	96	95.6
	rs2976396	143764001	**A:G**					95	93	96				97	95	95.2
	rs1045605	143764101	**G:A**	98	96	94	92	95	93	96				95	95	94.9
LY6K	rs1048831	143781858	**A:G**					46	50		27	29		23		35.0

T- tumor, N- normal bladder tissue sample; identical numbers for N and T samples mark paired samples

*****% frequency of first allele of each SNP in heterozygous samples as calculated based on RNA-seq reads; empty wells represent homozygous samples

† Average frequency of first allele of each SNP in heterozygous samples in normal and tumor bladder tissue samples

Supplementary Table 2. Allelic Expression Imbalance (AEI) in DNA and cDNA from tumor and adjacent normal bladder tissues.

Number	Sample ID	*Patient ID	Status	DNA†					cDNA†					paired *
				Allele C		Allele T		genotype	Allele C		Allele T		genotype	
				VIC:FAM	fold, T:N	VIC:FAM	fold, T:N		VIC:FAM	fold, T:N	VIC:FAM	fold, T:N		
1	BL_1383_T	17743	Normal	1.66		5.65		TC	0.54		2.41		TC	yes
2	BL_1015_T	31134	Normal	1.64		5.30		TC	0.63		1.59		TC	yes
3	BL_1024_T	32808	Normal	1.44		4.50		TC	0.72		1.56		TT	yes
4	BL_1480_T	39318	Normal	1.69		6.48		TC	0.78		2.03		TC	no
5	BL_1484_T	39319	Normal	1.54		5.74		TC	0.76		2.58		TC	yes
6	BL_1421_T	40194	Normal	1.64		4.48		TC	0.45		2.60		TT	yes
7	BL_1438_T	41885	Normal	1.43		5.88		TC	0.65		2.26		TC	yes
8	BL_1039_T	42824	Normal	1.51		4.48		TC	0.47		1.27		TT	yes
9	BL_1052_T	43791	Normal	1.56		5.66		TC	0.74		2.24		TC	no
10	BL_1447_T	44701	Normal	1.56		5.33		TC	0.59		2.27		TC	yes
11	BL_1455_T	46887	Normal	2.00		5.07		TC	0.71		1.16		TC	yes
12	BL_1457_T	46894	Normal	1.91		6.44		TC	0.49		1.75		TT	yes
13	BL_1512_T		Normal	1.54		4.72		TC	0.57		1.76		TT	no
14	BL_1516_T		Normal	1.72		5.67		TC	0.75		1.76		TC	no
1	BL_1382_T	17743	tumor	1.53	0.94	5.31	0.99	TC	0.48	0.77	2.98	1.53	TT	yes
2	BL_1014_T	31134	tumor	1.51	0.93	4.81	0.89	TC	0.54	0.86	2.29	1.18	TT	yes
3	BL_1023_T	32808	tumor	1.48	0.91	3.33	0.62	TC	0.81	1.29	2.48	1.28	TT	yes
4	BL_1483_T	39319	tumor	1.70	1.04	5.85	1.09	TC	0.41	0.65	2.02	1.04	TT	yes
5	BL_1420_T	40194	tumor	1.38	0.85	4.96	0.92	TC	0.39	0.62	3.28	1.69	TT	yes
6	BL_1439_T	41885	tumor	1.65	1.01	6.13	1.14	TC	0.75	1.19	2.73	1.41	TC	yes
7	BL_1038_T	42824	tumor	1.33	0.82	2.71	0.50	TC	0.46	0.73	2.22	1.14	TT	yes
8	BL_1448_T	44701	tumor	1.57	0.96	5.21	0.97	TC	0.53	0.84	2.75	1.42	TT	yes
9	BL_1454_T	46887	tumor	1.49	0.91	5.40	1.00	TC	0.38	0.60	2.57	1.33	TT	yes
10	BL_1456_T	46894	tumor	1.95	1.20	5.53	1.03	TC	0.23	0.36	3.78	1.95	TT	yes
				Allele C		Allele T			Allele C		Allele T			
		normal, average		1.63		5.38			0.63		1.94			
		tumor, average		1.55		5.14			0.49		2.88			
		average fold, T:N		0.95		0.95			0.78		1.48			
		T-test, T:N		**0.20**		**0.47**	**0.94**‡		**1.35E-02**		**1.26E-04**	**1.46E-06**‡		

* Identical patient IDs indicate paired normal-tumor samples

† Ratios are based on post-run allelic discrimination analysis using the same genotyping assay for rs2294008, VIC - C allele, FAM - T allele, VIC:FAM ratios represent relative amount of corresponding alleles in DNA and cDNA; fold, T:N is calculated for tumor samples as individual VIC:FAM ratios normalized by an a mean value in the group of normal samples.

‡ T-test for T:N ratio between C and T allele in DNA or cDNA

Quality of the allele-specific assay for rs2294008 was evaluated by a standard curve based on 9 serial dilution points of DNA samples with TT and CC genotypes and quantifying both alleles C (VIC) and T (FAM) fluorofores. There was a strong correlation between dilution and VIC and FAM signal intensities (r^2= 0.94).

Supplementary Table 3. Analysis of association between rs2294008 and PSCA IHC scores in 278 bladder tumor tissue samples, adjusting for relevant factors.

Association with PSCA IHC score	GLM multivariate model p-value*			Ordered logistic regression p-value†		
rs2294008	6.46E-11	3.71E-11	3.02E-11	1.03E-10	6.53E-11	5.15E-11
Age	0.54	0.49	0.52	0.52	0.47	0.50
Gender	0.57	0.45	0.41	0.49	0.42	0.37
Study site (ME, VT)	0.43	0.32	0.32	0.28	0.20	0.20
Ever smoke	0.54	0.57	0.52	0.50	0.53	0.49
Tumor stage	9.69E-03			6.43E-03		
Tumor grade		1.12E-03			8.92E-04	
Stage/grade combined			5.48E-04			4.20E-04

* Estimates from generalized linear models assuming additive genetic effect for rs2294008, adjusted for age, gender, study site, smoking status, tumor stage/grade when applicable.
† Estimates from ordered logistic regression models assuming additive genetic effect for rs2294008, adjusted for age, gender, study site, smoking status, tumor stage/grade when applicable.

Supplementary Table 4. Association between demographic factors, rs2294008 and PSCA IHC scores in 278 bladder tumor tissue samples.

		PSCA IHC scores					
		0, n=12 (4.3%)	**1, n=104 (37.4%)**	**2, n=115 (41.4%)**	**3, n=47 (16.9%)**		
		n (%)	**n (%)**	**n (%)**	**n (%)**	**Total (%)**	**p-value***
Age	**<50**	0 (0.0)	8 (34.8)	13 (56.5)	2 (8.7)	23 (8.3)	0.76
	50-54	1 (3.5)	12 (41.4)	10 (34.5)	6 (20.7)	29 (10.4)	
	55-59	4 (8.5)	22 (46.8)	14 (29.8)	7 (14.9)	47 (16.9)	
	60-64	1 (2.6)	15 (39.5)	15 (39.5)	7 (18.4)	38 (13.7)	
	65-69	2 (5.0)	17 (42.5)	15 (37.5)	6 (15.0)	40 (14.4)	
	70-74	1 (1.9)	16 (31.4)	25 (49.0)	9 (17.7)	51 (18.4)	
	75+	3 (6.0)	14 (28.0)	23 (46.0)	10 (20.0)	50 (18.0)	
Mean ± SD		65.6 ± 8.7	63.2 ± 9.6	64.4 ± 10.5	64.9 ± 9.2	64.1 ± 9.8	0.66
Male		10 (4.6)	83 (37.9)	86 (39.3)	40 (18.3)	219 (78.8)	0.49
Female		2 (3.4)	21 (35.6)	29 (49.2)	7 (11.9)	59 (21.2)	
Maine		8 (3.7)	78 (36.3)	95 (44.2)	34 (15.8)	215 (77.3)	0.30
Vermont		4 (6.4)	26 (41.3)	20 (31.8)	13 (20.6)	63 (22.7)	
Never smoker		1 (2.6)	18 (47.4)	12 (31.6)	7 (18.4)	38 (14.0)	0.52
Ever smoker		9 (3.9)	85 (36.5)	100 (42.9)	39 (16.7)	233 (86.0)	

Grade I		2 (1.5)	46 (34.1)	60 (44.4)	27 (20.0)	135 (48.6)	0.04
Grade II		3 (6.5)	15 (32.6)	17 (36.9)	11 (23.9)	46 (16.6)	
Grade III/IV		7 (7.2)	43 (44.3)	38 (39.2)	9 (9.3)	97 (34.9)	
Ta		4 (2.3)	61 (35.3)	75 (43.4)	33 (19.1)	173 (62.2)	0.60
T1		4 (6.7)	24 (40.0)	23 (38.3)	9 (15.0)	60 (21.6)	
T2		2 (8.7)	10 (43.5)	8 (34.8)	3 (13.0)	23 (8.3)	
T3/T4		2 (9.1)	9 (40.9)	9 (40.9)	2 (9.1)	22 (7.9)	
Low stage/grade		2 (1.6)	41 (32.0)	59 (46.1)	26 (20.3)	128 (46.0)	0.03
Moderate stage/grade		3 (5.9)	20 (39.2)	16 (31.4)	12 (23.5)	51 (18.4)	
High stage/grade		7 (7.1)	43 (43.4)	40 (40.4)	9 (9.1)	99 (35.6)	
rs2294008	**CC**	7 (8.9)	49 (62.8)	21 (26.9)	1 (1.3)	78 (28.1)	4.91E-09
	CT	4 (2.8)	41 (29.3)	68 (48.6)	27 (19.3)	140 (50.4)	
	TT	1 (1.7)	14 (23.3)	26 (43.3)	19 (31.7)	60 (21.6)	

*2-sided p-value of Chi-square or Student's T test