Fine mapping association study and functional analysis implicate a SNP in MSMB at 10q11 as a causal variant for prostate cancer risk

Bao-Li Chang1,2, Scott D. Cramer3, Fredrik Wiklund5, Sarah D. Isaacs7, Victoria L. Stevens8, Jielin Sun1,2, Shelly Smith1,2, Kristen Pruett1,2, Lina M. Romero3, Kathleen E. Wiley7, Seong-Tae Kim1,2, Yi Zhu1,2, Zheng Zhang1,2, Fang-Chi Hsu1,2,4, Aubrey R. Turner1,2, Jan Adolfsson6, Wennuan Liu1,2, Jin Woo Kim1,2, David Duggan9, John Carpten9, S. Lilly Zheng1,2, Carmen Rodriguez8, William B. Isaacs7, Henrik Gronberg5 and Jianfeng Xu1,2,3,*

1Center for Cancer Genomics, 2Center for Human Genomics, 3Department of Cancer Biology, 4Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA, 5Department of Medical Epidemiology and Biostatistics, 6Oncological Center, CLINTEC, Karolinska Institutet, Stockholm, Sweden, 7Johns Hopkins Medical Institutions, Baltimore, MD, USA, 8Department of Epidemiology and Surveillance Research, American Cancer Society, Atlanta, GA, USA and 9Translational Genomics Research Institute (TGen), Phoenix, AZ, USA

Received October 23, 2008; Revised January 8, 2009; Accepted January 15, 2009

A single nucleotide polymorphism (SNP) at 10q11 (rs10993994) in the 5' region of the MSMB gene was recently implicated in prostate cancer risk in two genome-wide association studies. To identify possible causal variants in the region, we genotyped 16 tagging SNPs and imputed 29 additional SNPs in ~65 kb genomic region at 10q11 in a Swedish population-based case–control study (CAncer of the Prostate in Sweden), including 2899 cases and 1722 controls. We found evidence for two independent loci, separated by a recombination hotspot, associated with prostate cancer risk. Among multiple significant SNPs at locus 1, the initial SNP rs10993994 was most significant. Importantly, using an MSMB promoter reporter assay, we showed that the risk allele of this SNP had only 13% of the promoter activity of the wild-type allele in a prostate cancer model, LNCaP cells. Curiously, the second, novel locus (locus 2) was within NCOA4 (also known as ARA70), which is known to enhance androgen receptor transcriptional activity in prostate cancer cells. However, its association was only weakly confirmed in one of the three additional study populations. The observations that rs10993994 is the strongest associated variant in the region and its risk allele has a major effect on the transcriptional activity of MSMB, a gene with previously described prostate cancer suppressor function, together suggest the T allele of rs10993994 as a potential causal variant at 10q11 that confers increased risk of prostate cancer.

INTRODUCTION

An association of prostate cancer risk with a SNP, rs10993994, at 10q11 was recently identified in two genome-wide association studies (GWAS) (1,2) and was confirmed in two additional studies (3,4). More importantly, the SNP is located at 57 bp upstream of the transcription start site (TSS) of the microseminoprotein beta gene (MSMB) that encodes PSP94, a primary constituent of semen and a proposed prostate cancer biomarker (5). Furthermore, this SNP was shown to affect promoter activities in human embryonic kidney cells (6). Although the statistical evidence of prostate cancer association and potential functional impact of this SNP are promising, it remains possible that the observed association may indirectly reflect...

*To whom correspondence should be addressed at: Center for Cancer Genomics, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA. Tel: +1 336 7137500; Fax: +1 336 713 7566; Email: jxu@wfubmc.edu

© The Author 2009. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
other sequence variants in the flanking region that are causally associated with prostate cancer risk.

In this study, we performed a fine mapping study to examine the association of all the known common sequence variants in the vicinity of rs10993994 at 10q11 with prostate cancer risk. We also examined the functional impact of implicated sequence variants on a prostate cancer cell line using promoter assays.

RESULTS
To explore whether additional SNPs in the regions flanking SNP rs10993994 are associated with prostate cancer risk, we genotyped 16 tagging SNPs in ~65 kb genomic region at 10q11 in the CAncer of the Prostate in Sweden (CAPS) study population. The distributions of genotypes for all the 16 SNPs were consistent with Hardy–Weinberg expectations in control subjects ($P > 0.05$). We also imputed 29 SNPs (with call rate >90%) in the region based on the genotyped SNPs using the computer program IMPUTE (7). Allele frequency differences between cases and controls in CAPS were compared for these 45 SNPs using a $\chi^2$ test (Supplementary Material, Table S1). Multiple SNPs in this ~65 kb region were significantly associated with prostate cancer risk (Fig. 1A, blue diamond). Specifically, many SNPs within a 25 kb region (51 194 000–51 219 000) were highly significant, including rs10993994 ($P = 0.001$); however, they were no longer significant after adjusting for rs10993994 (Fig. 1A, pink diamond). This group of SNPs was in strong linkage disequilibrium (LD) and located in one haplotype block, referred to here as locus 1 (Fig. 1B). Several SNPs that are immediately telomeric of locus 1 (51 219 000–51 234 000) were not significantly associated with prostate cancer risk. It is noted that there is a 7 kb genomic region to the telomeric side of rs10993994 where no SNP is cataloged in the HapMap database.

Interestingly, multiple SNPs that were further telomeric were again associated with prostate cancer risk (Fig. 1A, blue diamond). More importantly, they remained significantly associated with prostate cancer risk after adjusting for rs10993994 (Fig. 1A, pink diamond), suggesting they are independent from locus 1. These SNPs are in strong LD and located in a single haplotype block, and are termed locus 2. SNP rs10761581 (51 238 384) was the most significant ($P = 0.0005$) at locus 2 and the most significant in the entire ~65 kb region, even stronger than rs10993994 at locus 1. We estimated the recombination rate across the region among control subjects in CAPS using SequenceLDhot software (8) and found strong evidence for a recombination hotspot between the two loci at 51 219 000–51 238 000 ($P = 1.24 \times 10^{-15}$) (Fig. 1C). The recombination hotspot is also seen in the HapMap data (51 225 000–51 234 000, Release 21, Phases I and II). This recombination hotspot separates these two prostate cancer loci at 10q11. Prostate cancer risk-associated SNPs at locus 1 are in the 5’ untranslated region of the MSMB gene, and the SNP rs10993994 is 57 bp upstream of the TSS of MSMB. Prostate cancer risk-associated SNPs at locus 2, on the other hand, are within another gene, nuclear receptor coactivator 4 (NCOA4). NCOA4 is also known as androgen receptor coactivator (ARA70), which is known to enhance androgen receptor (AR) transcriptional activity in prostate cancer cells (9–11).

We also examined the association of these SNPs with prostate specific antigen (PSA) levels among control subjects (Fig. 1D). Similar to the results of Eeles et al. (2), we found the prostate cancer risk-associated allele of rs10993994 at locus 1 was strongly associated with higher PSA levels ($P = 0.001$). Surprisingly, although many genotyped or imputed SNPs at locus 1 are in strong LD with rs10993994, none of these SNP was significantly associated with PSA levels ($P > 0.05$). Similarly, none of the SNPs at locus 2 were significantly associated with PSA levels ($P > 0.05$).

When associations of SNPs at loci 1 and 2 were tested with clinicopathologic variables of prostate cancer cases (i.e. Gleason scores, TNM stages and pre-operative PSA), we did not observe any significant association ($P > 0.05$). However, we cannot exclude the possibility that these SNPs are weakly associated with the clinicopathologic variables. For example, the case–case study in CAPS can only detect with ≥80% power and at 5% significance level, if allele T of rs10993994 confers odds ratio (OR) ≥1.16 for aggressive prostate cancer over non-aggressive prostate cancer.

As a confirmation effort, we examined the novel locus 2, as well as the previously reported locus 1, in three additional study populations; The Johns Hopkins Hospital (JHH), Cancer Prevention Study-II (CPS-II) and Prostate, Lung, Colon and Ovarian (PLCO). One representative SNP in each locus (rs10993994 at locus 1 and rs10761581 at locus 2) was analyzed. A significant association was found for SNP rs10993994 at locus 1 in each population (Table 1). The overall $P$ of the allelic test in four populations was $7.7 \times 10^{-15}$ exceeding a genome-wide significance level of $2 \times 10^{-8}$ that accounts for multiple tests of ~2 million SNPs in the genome. It is noted that association results of rs10993994 at locus 1 in CAPS, PLCO and CPS-II have been previously published (1–4). For SNP rs10761581 at locus 2, the association was confirmed only in PLCO ($P = 0.02$) but not in JHH and CPS-II. The association of the SNP in PLCO remained significant after adjusting for rs10993994 at locus 1. Overall, the statistical evidence for locus 2 was weaker than that of locus 1, with a combined $P$ of 0.002 for an allelic test in four populations. The difference in allelic associations at locus 2 among these four study populations was marginally significant using a Breslow-Day test for homogeneity ($P = 0.05$), providing some evidence that the association at this locus is heterogeneous among studies.

The rs10993994 SNP is located at −57 bp 5’ to the TSS of the MSMB gene (Fig. 2A). To test the functional significance of the rs10993994 SNP, we cloned ~0.5 kb of the promoter sequence from individuals homozygous for either T or C at the rs10993994, and used these to drive luciferase expression in LNCaP cells. We found that the prostate cancer risk-associated allele T of the rs10993994 SNP had only 13% of the promoter activity compared with the C allele (Fig. 2B). Treatment with increasing concentrations of the synthetic androgen R1881 resulted in a dose-dependent increase in promoter activity of the C, but not the T allele of the SNP rs10993994. Because no other sequence variant within this ~0.5 kb region was found in these two clones, the different promoter activities are likely due to the variant rs10993994.
A similar promoter assay was performed for rs10761581; however, we did not observe differences in promoter activity between the two alleles (data not shown).

**DISCUSSION**

Using a fine mapping genetic association approach to follow-up a reported prostate cancer risk-associated SNP rs10993994 in a large population-based case-control study in Sweden (CAPS), we found two independent prostate cancer risk-associated loci at 10q11. Locus 1 included the originally reported prostate cancer risk-associated SNP rs10993994. The statistical evidence for prostate cancer association with this SNP was strong in the two original GWAS (1,2) and in two confirmation studies (3,4). An important contribution of our study is that we showed this SNP was the most significant SNP among all the known SNPs at locus 1 and the only significant SNP associated with PSA levels in the entire 65 kb at 10q11. Although these results provide critical positional information for further functional studies, we cannot rule out the possibility that other yet to be identified sequence variants in the region may be more strongly associated with prostate cancer risk.

In addition to the associated results, we found that the risk allele T of rs10993994 had greatly reduced MSMB promoter activity when compared with the C allele in LNCaP cell line model of prostate cancer. A previously reported functional reporter assay demonstrated that the T allele had 30% of the promoter activity of the C allele in HEK293T human embryonic kidney cells and minimal activity of either promoter in TE671 medulloblastoma cells (6), although this variant had not been evaluated in prostate cells. When searching for potential transcription factors that could bind the DNA sequence containing SNP rs10993994 using transcription element search system, we found that replacing the C allele with the T allele destroys a potential binding site (TGA C GT) for several ubiquitous transcription factors, such as cAMP response element binding protein and activating protein-1. This result corroborates our functional reporter assay that showed the T allele had much lower promoter activity compared with the C allele. In addition to the proximity of SNP rs10993994 to the TSS of the MSMB gene, we identified...
two potential androgen response element half sites (TGTTCT) in the MSMB promoter. One is located −113 to −118 bp 5’ to the TSS and is 23 bp upstream of SNP rs10993994 (Fig. 2A). There are also numerous putative GATA transcription factor binding sites (data not shown), which in other systems have been shown to influence tissue-specific AR signaling (12,13). MSMB has been described as a prostate cancer tumor suppressor gene on the basis that: (i) expression of MSMB typically decreases as prostatic epithelial cells become transformed and progress to advanced disease; and (ii) forced expression of PSP94 or exogenous administration of PSP94 or bioactive peptide fragments can induce prostate cancer apoptosis and suppress prostate cancer growth, invasion and metastasis (14–18). In fact, treatment with exogenous MSMB has been proposed as a therapeutic approach for metastatic prostate cancer. Thus, it is potentially highly significant that the risk allele T of rs10993994 that is associated with prostate cancer risk would be predicted to result in the production of lower amounts of this putative tumor suppressor gene in men harboring this variant. Although further studies will be necessary to more conclusively demonstrate that rs10993994 is the causal variant in this region and that its mechanism of action is to reduce the production of tumor suppressing MSMB, this possibility certainly seems consistent with the present data.

The prostate cancer risk association at locus 2 in our study, while novel, remains inconclusive. On the one hand, the statistical evidence for prostate cancer association at locus 2 was strong, and in fact, stronger than that of locus 1 in the CAPS study population. On the other hand, the association was only weakly confirmed in one of the three independent study populations. The fact that SNPs at locus 2 are within an interesting prostate cancer candidate gene (NCOA4, also known as ARA70) provides biological relevance for the genetic association. ARA70 is a ligand-dependent AR-associated protein that specifically enhances the transcriptional activity of AR in human prostate cancer cells in the presence of dihydropyrene or testosterone (19). As a potential facilitator of prostate cancer progression, ARA70-induced AR transactivation could result in the decreased apoptosis and increased cell proliferation in prostate cancer cells via a PSA-mediated mechanism as recently suggested by Niu et al. (20). Additionally, overexpression of an internally spliced 35 kDa ARA70 variant termed ARA70 beta promoted cellular invasion in an AR-independent manner (21). Further evaluation of this novel locus in additional study populations is needed.

If the second locus is confirmed in additional studies, this would be another example, following 8q24 (22–26) and 17q12 (27,28), where additional independent loci are subsequently discovered in the flanking regions of an initial locus identified from GWAS. The molecular mechanism for this phenomenon is unknown; however, it suggests that fine mapping studies in broad regions surrounding SNPs implicated from GWAS are needed for other prostate cancer risk-associated regions reported in recent GWAS.

The finding that rs10993994 was associated with PSA levels among control subjects in CAPS replicated the initial finding reported in the UK GWAS (2). In that study, the SNP was found to be significantly associated with PSA levels among 1646 control subjects (P = 8.5 × 10⁻⁶). The dual association of a SNP with both prostate cancer risk and PSA levels is the second confirmed example among the reported prostate cancer risk SNPs identified from GWAS. In the first example, a SNP in the 3’ of the KLK3 gene at 19q13 (rs2735839) was found to be significantly associated with prostate cancer risk and PSA levels in controls (2). The association of this SNP with PSA levels is biologically plausible because KLK3 (also known as the PSA gene) encodes for PSA protein and was consistently replicated in two other independent study populations (29,30). However, the association of the KLK3 SNP with prostate cancer risk remains controversial; negative (30) and positive (31) confirmations both have been reported among the independent study populations. It is suggested that the observed prostate cancer association of the KLK3 SNP may be due to PSA detection bias; men with the allele associated with higher PSA levels are more likely to be biopsied and diagnosed for prostate cancer than men who carry the allele associated with lower PSA level. Although other study designs are needed to dissect this PSA detection bias, results from these two examples suggest that a careful interpretation of prostate cancer risk association should be made for SNPs that are also associated with PSA levels in controls.

### Table 1. Association of prostate cancer risk with SNPs rs10993994 and rs10761581 (per study and combined)

<table>
<thead>
<tr>
<th>Risk allele</th>
<th>Genotype counts</th>
<th>Allele frequency</th>
<th>OR</th>
<th>95% LB</th>
<th>95% UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10993994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>CC   963</td>
<td>0.43</td>
<td>1.15</td>
<td>1.06</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>TC   1354</td>
<td>0.39</td>
<td>1.17</td>
<td>1.01</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>TT   546</td>
<td>1.5E⁻⁰³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JHH</td>
<td>T    442</td>
<td>0.47</td>
<td>1.24</td>
<td>1.10</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>TT   555</td>
<td>0.42</td>
<td>1.36</td>
<td>1.24</td>
<td>1.50</td>
</tr>
<tr>
<td>RS10761581</td>
<td>T    1481</td>
<td>0.7E⁻¹⁵</td>
<td>1.23</td>
<td>1.17</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>TT   1902</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPS</td>
<td>T    2343</td>
<td>0.42</td>
<td>1.17</td>
<td>1.07</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>TT   2389</td>
<td>0.42</td>
<td>0.99</td>
<td>0.86</td>
<td>1.15</td>
</tr>
<tr>
<td>JHH</td>
<td>G    507</td>
<td>0.45</td>
<td>1.15</td>
<td>1.02</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>TG   553</td>
<td>0.47</td>
<td>0.99</td>
<td>0.89</td>
<td>1.10</td>
</tr>
<tr>
<td>PLCO</td>
<td>T    464</td>
<td>0.43</td>
<td>0.06</td>
<td>1.23</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>TT   703</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RS10761581</td>
<td>T    1354</td>
<td>0.43</td>
<td>2.0E⁻⁰³</td>
<td>1.09</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>TT   1322</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aBased on allelic test assuming multiplicative model. The combined tests are based on M-H test.
bBreslow-Day test for homogeneity.
MATERIALS AND METHODS

Study subjects

Four study populations were included in the study. The first was a population-based prostate cancer case–control study in Sweden named CAPS and was used for the fine mapping study (32). Prostate cancer patients in CAPS were identified and recruited from regional cancer registries in Sweden. The inclusion criterion for case subjects was pathological- or cytological-verified adenocarcinoma of the prostate, diagnosed between July 2001 and October 2003. DNA samples from blood and TNM stage, Gleason grade (biopsy) and PSA levels at diagnosis were available for 2899 patients. Patients who met any of the following criteria were considered as having more aggressive disease: clinical stage T3/T4, N+, M+, Differential Grade III, Gleason score ≥8 or pre-operative serum PSA ≥50 ng/ml. Control subjects were recruited concurrently with case subjects. They were randomly selected from the Swedish Population Registry and matched according to the expected age distribution of cases (groups of 5-year intervals) and geographical region. DNA samples from blood were available for 1722 control subjects.

The second study population was a hospital-based case–control population at JHH (33). Prostate cancer cases were 1527 men of European descent (by self-report) who underwent radical prostatectomy for the treatment of prostate cancer at JHH from January 1, 1999 through December 31, 2006. Each tumor was graded using the Gleason scoring system and staged using the tumor–metastasis–node system. Patients who met any of the following criteria were considered as having more aggressive disease: pathologic Gleason score of 7 or higher, stage pT3 or higher, N+ or M1. Men undergoing screening for prostate cancer at JHH and in the Baltimore metropolitan area during the same time period were asked to participate as control subjects. A total of 482 men of European descent (by self-report) met our inclusion criteria as control subjects for this study: normal digital rectal examination, PSA levels ≤4.0 ng/ml and older than 55 years.

The third study population was selected from the American Cancer Society CPS-II Nutrition Cohort, a prospective study...
of cancer incidence (34). Approximately 184 000 US adults between the ages of 50 and 74 were enrolled in 1992 and sent follow-up questionnaires in 1997 and every 2 years afterwards. We identified Caucasian men who had been diagnosed with prostate cancer between 1992 and 2003 and had no previous history of cancer. Cancer status was verified through medical records, linkage with state cancer registries or death certificates. An equal number of controls were matched to the cases on age (± 6 months), race and date of blood collection (± 6 months) from men who were cancer-free at the time of cancer diagnosis of their matched case using risk set sampling. For SNP rs10993994, genotype data in 1773 cases and 1757 controls, as part of CGEMS study (1) were available for the analysis. For SNP rs10761581, only a subset of DNA samples (1429 cases and 1429 controls) was available for genotyping in this study.

The last study population was 1172 prostate cancer case subjects and 1157 control subjects who were selected from the PLCO Cancer Screening Trial and was the population for the first stage CGEMS prostate cancer GWAS (1). Individual genotype data were obtained through an approved data request application. For both CPS-II and PLCO, patients with Gleason score ≥7 or Stage ≥ III were considered as aggressive prostate cancer.

**Region of interest, tagging SNPs and SNP genotyping**

We identified ~64 kb region of interest for the fine mapping study (51 194 000–51 259 000, Build 35) based on (i) the association results at 10q11 from the CGEMS GWAS where multiple SNPs were significantly associated with prostate cancer risk \((P < 0.05)\) and (ii) the known genes \((MSMB)\) and \((NCOA4)\) in the implicated region. A total of 16 tagging SNPs were identified to capture \((r^2 > 0.8)\) all the SNPs with minor allele frequencies of 1% higher in the region of interest based on the HapMap Phase II data. These SNPs were genotyped in CAPS. Polymerase chain reaction (PCR) and extension primers for these SNPs were designed using the MassARRAY Assay Design 3.0 software. PCR and extension reactions were performed according to the manufacturer’s instructions, and extension product sizes were determined by mass spectrometry using the Sequonem iPLEX system. Duplicated and water samples, to which the technician was blind, were included in each 96-well plate as PCR negative controls. The average genotype call rate for these SNPs was >98%, and the average concordance rate was 99.7% among 100 duplicated quality control samples.

**Statistical methods**

A Hardy–Weinberg equilibrium test was performed using the Fisher’s exact test. Haplotype blocks were estimated using a computer program Haploview (35), and a default Gabriel method (36) was used to define a haplotype block; i.e. a region in which all (or nearly all) pairs of markers are in ‘strong LD’, which is consistent with no historical recombination. SequenceLDhot was used to determine recombination hotspots (8). SequenceLDhot considers a grid of putative hotspot positions, and for each putative hotspot it calculates a likelihood ratio statistic for the presence of a hotspot.

Haplotype and background recombination rates generated from PHASE (version 2.1) were used as input files. We assumed the putative hotspots have a width of 2 kb and the program considers a new hotspot every 1 kb.

We imputed all of the known SNPs in the genome based on the genotyped SNPs and haplotype information in the HapMap Phase II data (CEU) using a computer program, IMPUTE (7). A posterior probability of 0.9 was used as a threshold to call genotypes.

Allele frequency differences between case patients and control subjects were tested for each SNP, using a \(\chi^2\) test with one degree of freedom. The allelic OR and 95% confidence interval were estimated based on a multiplicative model. Results from multiple case–control populations were combined using a Mantel–Haenszel model in which the populations were allowed to have different population frequencies for alleles but were assumed to have a common OR. The homogeneity of ORs among different study populations was tested using Breslow-Day \(\chi^2\) test. Independence of prostate cancer associations with SNPs at two loci at 10q11 was tested by including both SNPs (assuming an additive model at each SNP) in a logistic regression model among four populations and adjusted for study population and age. Multiplicative interactions between two SNPs were tested by including both SNPs (assuming a general model) and an interaction term (product of two main effects) in a logistic regression model.

We also tested the association of SNPs with PSA levels in controls by assuming an additive model and adjusting for age using a multiple regression analysis. PSA levels were logarithm transformed to best approximate the assumption of normality.

Associations of SNPs with aggressiveness of prostate cancer and TNM stages were tested among cases only using a \(\chi^2\) test of \(2 \times N\) table. A trend test was used to assess the proportion of prostate cancer-associated genotypes with each increasing Gleason score, from ≤4 to 10. All reported \(P\)-values were based on a two-sided test.

**Vector construction and luciferase assay**

PCR fragments containing either C or T allele of rs10993994 were amplified from genomic DNAs isolated from homozygous C or homozygous T carriers of SNP rs10993994 using the following primers: forward primer \((5’-AGCTCTACGC GTGAAGGTCCAGCAATTCAGC-3’)\), reverse primer \((5’-AGCTCTCAGGACGAGATCCTTATAGACC-3’)\). PCRs were carried out as follows: initial denaturation at 95°C for 3 min was followed by 35 cycles of denaturing at 94°C for 20 min, annealing at 58°C for 20 min and extension at 69°C for 40 min. The PCR products were then cloned into pGL3-Basic vector (Promega, Madison, WI, USA) between \(MluI\) and \(XhoI\) (Promega) sites. The sequences of the cloned PCR fragments containing the expected C or T allele of rs10993994 were verified by sequencing.

LNCaP human prostate cancer cells (passages 21–30) were transfected as previously described (37) with the following modifications. Briefly, LNCaP cells were plated on 6-well plates at \(3 \times 10^5\) cells per well using RPMI 1640 containing 10% fetal bovine serum (FBS) and allowed to grow for 2 days prior to transfection \((\approx 70\%\) confluence). Transfection experiments were performed by using lipofectamine 2000.
(Invitrogen, Carlsbad, CA, USA). All plasmids were purified using EndoFree -MAXIPREP (Qiagen, Valencia, CA, USA). Two hours prior to transfection, the medium was changed to naked Dulbecco’s modified Eagle’s medium (2 ml). For each well, 100 μl of transfection cocktail was added directly to the well. Transfection cocktail was made by mixing 50 μl of a 1:10 dilution of Lipofectamine 2000:Opti-MEM medium (Invitrogen) with 50 μl of Opti-MEM containing 2 μg of experimental pGL-3 vector (C or T vectors above) and 0.1 μg of renilla vector (pRLTK, Promega). The resulting mix was incubated with gentle shaking for 45 min at room temperature prior to addition to cells. Each transfection was performed in triplicate. After 24 h, the medium was changed to experimental medium (RPMI-1640, plus 10% FBS), containing the indicated doses of the synthetic androgen R1881 (Methyltrienolone, NEN Life Sciences, Boston, MA, USA) or vehicle control (0.1% ethanol). Twenty-four hours after treatment with R1881, cell lysates were made using the passive lysis buffer provided in the Dual Luciferase Kit (Promega), and firefly luciferase and Renilla luciferase activities were measuring in the reporter microplate lumino-

REFERENCES


SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

The authors thank all the study subjects who participated in the CAPS study and urologists who included their patients in the CAPS study. The authors acknowledge the contribution of multiple physicians and researchers in designing studies and recruiting study subjects, including Dr Hans-Olov Adami (for CAPS) and Drs Bruce J. Trock, Alan W. Partin and Patrick C. Walsh (for JHH).

The authors also thank the National Cancer Institute Cancer Genetic Markers of Susceptibility Initiative (CGEMS) for making the data available publicly.

Conflict of Interest statement. None declared.

FUNDING

The work was supported by National Cancer Institute (grant numbers CA129684, CA105055, CA106523 and CA95052 to J.X.; CA112517 and CA58236 to W.B.I.); Department of Defense (grant number PC051264 to J.X.); Swedish Cancer Society (Cancerfonden) to H.G. and Swedish Academy of Sciences (Vetenskapsrådet) to H.G. The support of Kevin P. Jaffe to W.B.I. is gratefully acknowledged.


