Carcinogenesis vol.28 no.5 pp.1032–1039, 2007
doi:10.1093/carcin/bgl236
Advance Access publication December 6, 2006

PSA/KLK3 AREI promoter polymorphism alters androgen receptor binding and is associated with prostate cancer susceptibility

John Lai[1,3,†], Mary-Anne Kedda[2,3,*,†], Kimberly Hinze[2,3], Robert L.G.Smith[1,2,3], John Yaxley[4], Amanda B.Spurdle[5], C.Phillip Morris[1,3], Jonathan Harris[1,3] and Judith A.Clements[1,3]

[1]Cluster for Molecular Biotechnology, School of Life Sciences and Science Research Centre, [2]Centre for Health Research, School of Public Health, [3]Institute of Health and Biomedical Innovation, Queensland University of Technology, Victoria Park Road, Kelvin Grove, Brisbane, QLD 4059, Australia, [4]Brisbane Private Hospital, Brisbane, QLD 4000, Australia and [5]Molecular Cancer Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, QLD 4029, Australia

*To whom correspondence should be addressed. Tel: +61 0 7 3864 9674; Fax: +61 0 7 3864 3369;
Email: m.kedda@qut.edu.au

The proximal promoter of the kallikrein-related peptidase 3 gene (*KLK3/PSA*) contains a single-nucleotide polymorphism (G-158A) located within the second canonical half-site for the prostate-specific antigen (*PSA*) androgen response element 1 (AREI). Previous studies suggest that this polymorphism may be associated with higher PSA levels and increase prostate cancer risk. We have investigated the potential functional significance of this polymorphism and its association with prostate cancer susceptibility by genotyping the G-158A polymorphism in 209 men diagnosed with prostate cancer and 223 healthy control men in an Australian Caucasian population. Functional analyses of *PSA* AREI demonstrated that the A allele increased binding of AREI to the androgen receptor, as well as increasing transcriptional response to androgens. Association studies of the G-158A polymorphism demonstrated that men with an A/A genotype had a 3-fold increased risk for developing prostate cancer [95% confidence intervals (CIs) = 1.36–6.52] and men with an A/G genotype had a 2.4-fold increased risk (95% CIs = 1.23–4.81). Under a dominant model, the A allele conferred a 2.6-fold increased risk for prostate cancer (95% CIs = 1.37–4.96, *P* = 0.004). Taken together with the finding that the G-158A polymorphism is associated with an increased risk of prostate cancer in Australian men, our functional data suggest that the presence of the A allele in AREI may, in part, account for the altered *PSA* regulation seen in prostate cancer.

Introduction

Prostate-specific antigen (PSA, kallikrein-related peptidase 3) is a serine protease that is part of the kallikrein superfamily (1,2), produced predominantly by the prostate and primarily by secretory luminal epithelial cells therein (3). Serum PSA is the most commonly used clinical biomarker for prostate cancer. In the prostate, *PSA* is regulated by androgens and the cellular effects of androgens are mediated by the androgen receptor (AR). Upon androgen binding to the AR, a sequence of cellular events occurs that ultimately results in the translocation of the AR into the nucleus, where it binds to androgen response elements (AREs) in the promoters of target genes to initiate transcription. These AREs consist of palindromic hexameric repeats separated by a three-nucleotide spacer (GGTACAnnnTGTTCT) (4). Three AREs have been described in the promoter of the *PSA* gene, two are found within the proximal promoter at −170 (AREI) and −394 (AREII) and the third ARE is located in the distal enhancer region at −4200 (AREIII) (5–7).

Cramer *et al.* (8) have shown that a number of polymorphisms in the *PSA* gene are associated with changes in serum PSA levels and a recent editorial has emphasized the importance of these polymorphisms as predictive biomarkers (9). The *PSA* AREI harbours a polymorphism (rs266882) that results in a substitution of a guanine by an adenine 158 bases upstream of the transcription start site (G-158A) (10). This polymorphism is located in the second canonical half-site of AREI, and potentially affects transcriptional control of AR-regulated expression. Indeed, this single-nucleotide polymorphism (SNP) has been associated with increased serum PSA levels (11–14). It was recently reported that there was no difference in transcriptional activity between the two AREI alleles in response to androgens in both LNCaP and PC-3 cells co-transfected with wild-type AR (15); however, it is possible that subtle differential interactions between the two AREI alleles and the AR were not measurable in that study (15).

Given the potential importance of the *PSA* G-158A SNP, a number of studies have investigated the association between this SNP and prostate cancer susceptibility in different populations. An initial association study found a 3-fold increased risk for developing prostate cancer in men with the GG genotype and, furthermore, this risk was compounded to 5-fold when associated with short AR alleles (16). However, subsequent studies have been conflicting, showing an association between increased risk of prostate cancer with either the A (14,17,18) or G (19,20) allele, whereas others have found no significant association with either allele (21–23). The G-158A SNP has also been reported to be associated with prostate tumour volume, stage/grade of disease and circulating tumour cells (11–14,16,20).

Given the importance of the AR-signalling axis in prostate cancer aetiology, and the observation that serum PSA is frequently elevated in men with prostate cancer, we used a number of *in vitro* assays to elucidate the functional significance of this polymorphism. We also conducted a case–control study in an Australian Caucasian population, in order to determine the association between the *PSA* G-158A SNP and risk of prostate cancer in this population.

Abbreviations: AR, androgen receptor; ARE, androgen response element; BPH, benign prostatic hyperplasia; CI, confidence interval; DBD, DNA-binding domain; DHT, dihydrotestosterone; EMSA, electrophoretic mobility shift assay; OR, odds ratio; SNP, single-nucleotide polymorphism; PCR, polymerase chain reaction; PSA, prostate-specific antigen; TNM, tumor, node, metastasis.

†These authors contribute equally to this work.

Materials and methods

Tissue culture and reagents

LNCaP and 22Rv1 prostate cancer cell lines were obtained from American Type Tissue Culture Collection (Manassas, VA). All cells were initially maintained in RPMI 1640 supplemented with 10% fetal calf serum (Invitrogen, Carlsbad, CA) and 50 U/ml penicillin G and 50 μg/ml streptomycin (Roche, Basel, Switzerland), and grown in a 37°C incubator in 5% CO_2. At ∼70% confluency, cells were serum-starved for 48 h in 2% charcoal-stripped fetal calf serum (Hyclone, South Logan, UT), prior to the addition of androgens for androgen regulation experiments.

Purification of AR–DNA-binding domain

A fragment encoding 117 amino acids of the AR–DNA-binding domain (DBD) and the ligand-binding domain was amplified from T-47D breast cancer cDNA

by standard polymerase chain reaction (PCR) using primers ARDBD-F (5′-ctctctggatccgccagggaccatgttttg-3′) and ARDBD-R (5′-ctactagaattccttctgggttgtctcctc-3′). Primers were designed with BamHI and EcoRI restriction sequence overhangs to facilitate subcloning into the pGEX-2T multiple-cloning site (Amersham Biosciences, Uppsala, Sweden) prior to transformation into *Escherichia coli* JM109 cells (Stratagene, La Jolla, CA). Fusion protein, AR-DBD/glutathione *S*-transferase, expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactoside (Promega, Madison, WI) and the fusion protein was purified by batch chromatography on glutathione-sepharose 4B (Amersham, Piscataway, NJ). Purification and protein size was assessed by sodium dodecylsulphate–polyacrylamide gel electrophoresis and quantified by Coomassie Brilliant Blue protein staining (Bio-Rad, Hercules, CA).

Electrophoretic mobility shift assay

Oligonucleotide sequences used in electrophoretic mobility shift assay (EMSA) and limited proteolysis assays were designated PSA-AREI-G (sense: 5′-tcgacttgcAGAACAgcaAGT**G**CTagctg-3′ and anti-sense: 5′-agctcagctAG**C**ACTtgcTGTTCTgcaag-3′; where the ARE is underlined, base changes are indicated in bold and the hexameric half-sites are represented by capital letters) or PSA-AREI-A (sense: 5′-tcgacttgcAGAACAgcaAGT**A**CTagctg-3′ and anti-sense: 5′-agctcagctAG**T**ACTtgcTGTTCTgcaag-3′). DNA oligonucleotides were labelled using the biotin 3′ end DNA labelling kit (Pierce Biotechnology, Rockford, IL) and annealed to their complementary strand. EMSA was carried out using the LightShift Chemiluminescent EMSA Kit (Pierce Biotechnology) according to the manufacturer's instructions. Bands were visualized and quantitated using the Bio-Rad GS-690 imaging densitometre and the Bio-Rad Multi-Analyst software (Version 1.1, Bio-Rad). Binding experiments with glutathione *S*-transferase and PSA-AREI$_A$ were carried out to ensure that the AR–DBD and not the glutathione *S*-transferase moiety was complexing with the *PSA* AREs (data not shown). Experiments were conducted in triplicate and data are represented as the mean ± SE from three independent experiments. The K_d values were determined using allosteric Hill kinetics and the corresponding binding curves were determined by the SigmaPlot statistical software (Version 9.0, Systat Software).

Limited proteolysis assay

EMSA was carried out as above, with the addition of 10 μM $CaCl_2$ to stabilize trypsin digestion, and performed on 5 μg of nuclear extract from LNCaP and 22Rv1 cells treated with 10 nM R1881 for 12 h. Following 20 min incubation for DNA–protein complex formation, 0, 10, 20 and 40 ng of trypsin (Worthington Biochemical Corporation, Templestowe, Australia) was added for an additional 1 h at room temperature. Samples were then electrophoresed on 8% non-denaturing acrylamide gels, followed by DNA transfer and visualization as described above.

Luciferase promoter constructs and reporter assays

PSA luciferase promoter constructs for the G and A alleles were generated using oligonucleotides (PSA AREIX3-G sense: 5′-cttgcAGAACAgcaAGT**G**CTagctctctaattgcAGAACAgcaAGT**G**CTagctctctaattgcAGAACAgcaAGT**G**CTagcta-3′ and anti-sense: 5′-gatctagctAG**C**ACTtgcTGTTCTgcaattagagagctAG**C**ACTtgcTGTTCTgcaattagagagctAG**C**ACTtgcTGTTCTgcaaggtacc-3′ and PSA AREIX3-A sense: 5′-cttgcAGAACAgcaAGT**A**CTagctctctaattgcAGAACAgcaAGT**A**CTagctctctaattgcAGAACAgcaAGT**A**CTagcta-3′ and anti-sense: 5′-gatctagctAG**T**ACTtgcTGTTCTgcaattagagagctAG**T**ACTtgcTGTTCTgcaattagagagctAG**T**ACTtgcTGTTCTgcaaggtacc-3′; where the ARE is underlined, base changes are indicated in bold and the hexameric half-sites are represented by capital letters) consisting of three tandem copies of *PSA* AREI and its native flanking sequences (Proligo, Lismore, Australia), and designated PSA-AREIX3-G-Luc and PSA-AREIX3-A-Luc, respectively. The inserts were designed with BglII and KpnI restriction site overhangs to facilitate orientation-specific ligation into the multiple-cloning site of the luciferase pGL3 promoter vector (Promega). A 'full-length' *PSA* promoter construct (PSA-5.8-A-Luc) was generated using primers targeting ∼5.8 kb of the *PSA* promoter (PSA 5.8-A-F: 5′-cacctcgagttcttttcccggtgacatcg-3′ and PSA 5.8-A-R: 5′-tcacggacagggtgaggaag-3′), amplified from LNCaP DNA. The PCR amplicon was cloned into pGEM-T and digested with HindIII and XhoI prior to ligation into the promoter-less pGL3-basic vector (Promega). All promoter constructs were verified by restriction mapping and direct sequencing.

Luciferase reporter assays were carried out in both LNCaP and 22RvI cells. Cells were seeded in 24-well plates at a density of $\sim 1 \times 10^5$ cells per well and culture medium was changed to phenol red-free RPMI 1640 and 2% charcoal-stripped fetal calf serum for 48 h prior to transfection. Transient transfection was carried out with PSA-AREIX3-G-Luc, PSA-AREIX3-A-Luc, pGL3 promoter or PSA-5.8-A-Luc with Opti-MEM®-I reduced serum medium (Invitrogen), and 2 μl of lipofectamine™ 2000 (Invitrogen) per well. Cells were co-transfected with the Renilla luciferase construct (phRL-TK, Promega) to serve as a transfection efficiency control. Luminescence was measured by the Dual-Luciferase Reporter Assay System (Promega). Androgen responsiveness of the constructs was assessed by treating transfected cells with either 10 nM of dihydrotestosterone (DHT) or 10 nM of the synthetic androgen R1881. Statistical significance was assessed using the Student's *t*-test from three independent data sets that were each carried out in quadruplicate.

Molecular dynamic simulation/modelling

All molecular modelling manipulations were carried out within the YASARA Dynamics Program (Version 6.2.4) (24). The starting point was the previously published structure of the AR–DBD, binding to a canonical ARE PDBid 1R4I (25). The sequence of the canonical ARE was mutated *in silico*, so that it accorded with either the −158A or G sequences. All water molecules resolved in the starting structure were removed, hydrogens were added where missing and the following missing lysines were reconstructed: A protomer Lys 573 and 575 and B protomer 540, 573, 575 and 588. Molecular dynamic simulation was carried out with a cut-off of 7.6 Å to delimit non-bonding interactions, and a particle Ewald mesh approximation was used for longer range electrostatic forces. Simulations were carried out in a cell with periodic boundaries that extended 10 Å beyond the target structure in all three axes. This was filled with water molecules, which were then subject to simulated annealing energy minimization. Counter ions (six sodium ions) were then added to neutralize the ensemble, pK_a values for all side chains were calculated and protons placed according to ionization status. Following neutralization, resulting conformational stress in the protein was released by energy minimization. Simulations were then conducted at 298 K, at constant pressure, with molecular snapshots being taken every 10 ps for a total of 500 ps. Snapshots for every 50 ps were then overlaid and differences in hydrogen bonding and residues displacements between the −158A and −158G structures were tabulated and identified. Calculation of root mean square deviations between the starting structure and the simulated complex showed good concordance. Average root mean square deviation for all atoms of the structure was 1.13 Å (structure not shown), indicating that the simulation procedure did not significantly change the arrangement of the naturally folded protein structure.

Patient samples

Men recently diagnosed with prostate cancer ($n = 209$) and healthy male blood donor controls ($n = 223$) participated in the association study. Prostate cancer was defined as clinically diagnosed, and pathologically confirmed, cancer of the prostate, following presentation with an abnormal serum PSA and/or lower urinary tract symptoms. The healthy controls had no personal history of any cancer. Men with prostate cancer were recruited within 2 years of diagnosis, through physician referrals from Brisbane Private Hospital ($n = 191$), Royal Brisbane Hospital ($n = 4$) and Princess Alexandra Hospital ($n = 14$) in Brisbane, Queensland. Healthy male blood donors were recruited through the Australian Red Cross Blood Services in Brisbane, Queensland. All participants were unrelated Caucasian men between 18 and 91 years of age, and all subjects gave written informed consent and were interviewed by one of the investigators. Comprehensive epidemiological questionnaires and ∼8 ml of blood were obtained for each participant and, in addition, detailed clinical data (including serum PSA levels at diagnosis, tumor, node, metastasis (TNM) staging and Gleason score) were obtained for cases. All samples and data from the controls were deidentified. The study protocol was approved by the Human Research Ethics Committees of Queensland University of Technology, Queensland Institute of Medical Research, the Mater Hospital (for Brisbane Private Hospital) and the Royal Brisbane and Princess Alexandra Hospitals, Queensland.

DNA extraction and genotyping

Genomic DNA was extracted from peripheral blood leukocytes using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions, or using the salt extraction method described by Miller *et al.* (26).

Genotypic analysis of the PSA G-158A SNP was carried out using PCR–restriction fragment length polymorphism analysis as described previously (27). Briefly, 25 μl reactions containing 50 ng genomic DNA, 20 pmol of each primer (PSA-F and PSA-R), 1× PCR buffer, 0.75 mM $MgCl_2$, 200 mM dNTP and1 U *Taq* polymerase (Platinum *Taq*, Invitrogen) were cycled at 95°C for 2 min, followed by 28 cycles of 94°C for 30 s, 59–64°C for 30 s, 72°C for 30 s and a final cycle of 72°C for 8 min. Amplicons were digested at 37°C for 4–8 h with 10 U restriction endonuclease NheI (New England Biolabs, Beverley, MA) and electrophoresed on 2.5% ethidium bromide-stained agarose gels. Genotypes were verified by repeating PCR–restriction fragment length polymorphism on 20 random samples.

Statistical analysis

Genotype and allele frequencies for the PSA G-158A polymorphism were calculated for the patient and healthy control groups. Comparisons of allele and genotype distribution and their association with prostate cancer susceptibility, socio-demographics and clinical data were performed under co-dominant and dominant models, using chi-square test and multivariate logistic regression analysis, in SPSS for Windows (Release 13.0, 2004. Chicago: SPSS) Hardy–Weinberg equilibrium analysis for each group was evaluated using Ottutil Utility Program for Analysis of Genetic Linkage (J. Ott, 1998).

Results

The PSA AREI alleles bind with different affinity to the AR

EMSA experiments were carried out to assess whether there was any difference in AR-binding affinity between the two *PSA* AREI alleles, using a fragment comprising 117 amino acids of the AR–DBD and ligand-binding domain (Figure 1A). Incubation of PSA-AREI-G and PSA-AREI-A with 0.03–1.5 μg of AR–DBD resulted in the formation of a distinct high-molecular weight shift for both alleles (Figure 1B), and PSA-AREI-A bound with greater affinity compared with PSA-AREI-G across all concentrations of AR–DBD. The K_d binding curves for PSA-AREI-G and PSA-AREI-A are presented in Figure 1C and show a 2-fold difference in binding affinities between the alleles (0.6 ± 0.15 μM and 0.3 ± 0.05 μM, respectively).

The PSA AREI alleles mediate differential binding of endogenous nuclear receptor complexes and are cell specific

Limited proteolysis assays were carried out to assess whether the *PSA* AREI alleles mediated differential co-activator complex stability and/or altered endogenous AR conformation *in vitro*. When incubated with nuclear extracts from both LNCaP and 22Rv1 cells treated with R1881, both AREI alleles showed multiple shifts, with only two shifts in common (Figure 2). Differences in shift intensities between the alleles for numerous protein–AREI complexes suggest that these alleles bind with different affinity to the receptor complexes. Furthermore, proteolytic protection of protein–AREI complexes for the A allele was observed in samples that were treated with trypsin and indicates marked differences in proteolytic sensitivity between protein–AREI-G/A complexes in both LNCaP and 22Rv1 cells.

The PSA AREI A allele is more transcriptionally responsive to androgens

Reporter assays were carried out with constructs containing three tandem copies of either PSA AREI allele (PSA-ARE1X3-G-Luc or PSA-ARE1X3-A-Luc) (Figure 3A), and androgen-induced transactivation assays were carried out with DHT or R1881 in both LNCaP and 22Rv1 cells, in order to determine whether there were allelic differences in the induction of AR activity. The full-length promoter construct, PSA-5.8-A-Luc, was androgen responsive in both cell lines when compared with vehicle control (Figure 3B and C), and there was <1-fold difference in pGL3 promoter activity in both DHT- and R1881-treated LNCaP and 22Rv1 cells, confirming that the vector backbone was not androgen responsive (Figure 3B and C).

The *PSA* AREI-A allele was more responsive to both DHT and R1881 in both LNCaP and 22Rv1 cells, compared with the PSA AREI-G allele. Although not statistically significant, there was a 1.35-fold (DHT) and 6-fold (R1881) higher transcriptional response for PSA-AREIX3-A-Luc compared with PSA-AREIX3-G-Luc in LNCaP cells (Figure 3B). In 22Rv1 cells, PSA-AREIX3-A-Luc had a 1.6- and 1.4-fold higher transcriptional activity over PSA-AREIX3-G-Luc in response to R1881 and DHT, respectively (Figure 3C). The difference in androgen regulation between PSA-AREIX3-A-Luc and PSA-AREIX3-G-Luc was statistically significant for the DHT-treated 22Rv1 cells ($P = 0.04$).

The G-158A polymorphism introduces two extra hydrogen bonds with the AR–DBD

A series of *in silico* simulations of AR–DBD binding to the polymorphic ARE were conducted to achieve a structural understanding of the differences in K_d for the two alleles. The simulation procedure was conducted with the −158A/G sequences substituted for the canonical ARE and root mean square deviation calculations on the resulting complexes showed substantially higher average deviations of 1.79 Å for the A allele and 1.63 Å for the G allele compared with

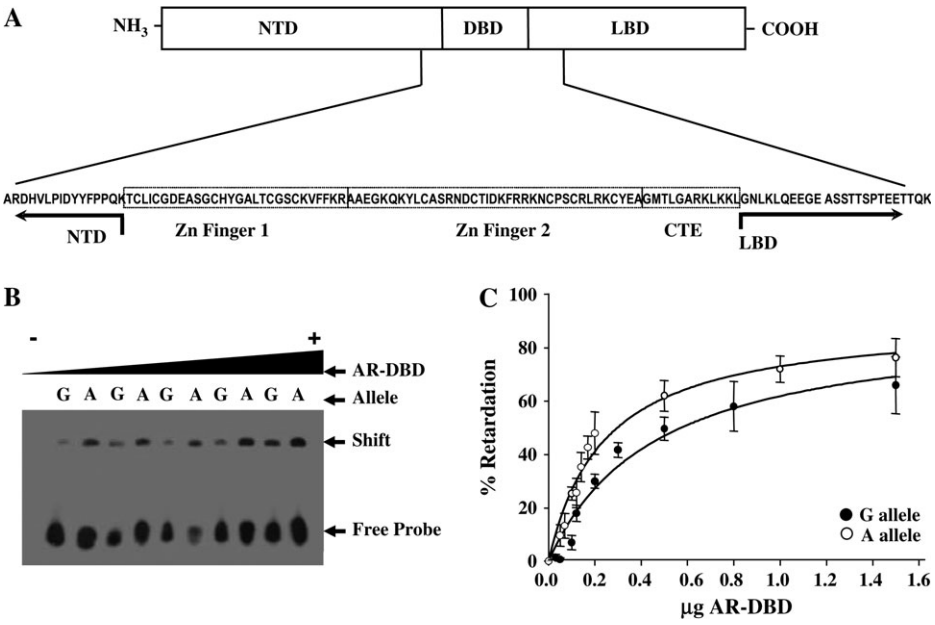

Fig. 1. EMSA analysis for differences in binding affinity conferred by the G-158A SNP with the AR–DBD. (**A**) Schematic of human AR–DBD used in EMSA and limited proteolysis experiments. The purified AR–DBD comprises residues of the N-terminal domain (NTD), the two zinc finger motifs (Zinc Finger 1/2) in the DBD and a region of the ligand-binding domain (LBD) that includes the C-terminal extension (CTE). (**B**) Band shifts of PSA-AREI-G (lanes G) and PSA-AREI-A (lanes A) with increasing amounts (0.03–1.5 μg) of AR–DBD. (**C**) Binding curves plotting the percentage of complexed DNA probe (% retardation) for PSA-AREI-A (open circles) and PSA-AREI-G (closed circles) with 0.03–1.5 μg of AR–DBD. Data are represented as the mean ± SE from three independent experiments.

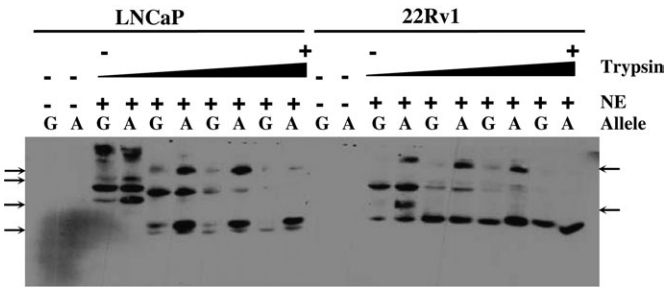

Fig. 2. Limited proteolysis analysis assessing differential AR conformation and/or complex stability conferred by the *PSA* G-158A SNP in LNCaP and 22Rv1 cells. Nuclear fractions were extracted from cells that were treated with 10 nM R1881 for 12 h. After formation of receptor–DNA complexes by incubating nuclear extracts (NE) with biotin-labelled PSA AREI-G or PSA AREI-A, samples were then digested with increasing amounts of trypsin (0—40 ng) for 1h at room temperature. Arrows indicate shifts that have different intensity between the two alleles.

the starting template structure. The majority of this displacement occurred in the B protomer (Figure 4A), which contacts the polymorphic bases in the ARE. Comparative analysis of protomer B for both the −158A and −158G complexes showed that maximum displacement occurred at residue ARG_{568}. This displacement corresponds with the formation of two extra hydrogen bonds with base −160G of the oligonucleotide (Figure 4B). Significantly, the equivalent residue in the A protomer shows very little displacement when the two model complexes are compared. The binding energy of the DNA–protein complex for both the −158A and −158G alleles was calculated as 40.0 and 36.9 kJ, respectively, which was comparable with the dissociation constants measured previously by EMSA analysis. This provided validation of the simulation approach used to analyse the structural features responsible for differences in binding between the alleles.

The A allele in PSA AREI is associated with increased prostate cancer risk in Australian men

In total, 209 men recently diagnosed with prostate cancer and 223 healthy male blood donor controls participated in this study. Table I illustrates a number of the socio-demographic and clinical characteristics of the populations studied. Frequencies for the *PSA* G-158A polymorphism in the study groups are summarized in Table II and both cases and controls were in Hardy–Weinberg equilibrium (P = 0.67 and P = 0.31, respectively). Under a co-dominant model (adjusted for age), both the A/G and A/A genotypes were associated with an increased risk of developing prostate cancer (Table II): men with the A/G genotype had a 2.4-fold increased risk [95% confidence intervals (CIs) = 1.23–4.81] and the A/A genotype was associated with a 3-fold increased risk (95% CIs = 1.39–6.52), with significant evidence for trend (P = 0.012). Under a dominant model (adjusted for age), the A allele was associated with a 2.6-fold increased risk of developing prostate cancer (95% CIs = 1.37–4.96, P = 0.004). There was no allelic or genotypic association between the *PSA* G-158A polymorphism and various clinical markers in men with prostate cancer, including serum PSA at diagnosis (P = 0.34), TNM staging (P = 0.64) or Gleason score (P = 0.86). Although all analyses were adjusted for age, the age distributions of the cases and controls were significantly different ($P < 0.0001$). An additional subgroup analysis, therefore, was performed to determine whether a better age-matched control group (i.e. controls aged 50 years and over, n = 115, median age = 60 years) gave similar results. In fact, although the crude odds ratios (ORs) were somewhat higher for the age-restricted group as controls on both a co-dominant model [OR_{AG} = 2.41 (95% CIs = 1.38–4.20) and OR_{AA} = 3.26 (95% CIs = 1.69–6.31)] and a dominant model [crude $OR_{AA/AG}$ = 2.66 (95% CIs = 1.58–4.50), $P < 0.0001$], the results of the age-adjusted analyses were very close to those for the overall group

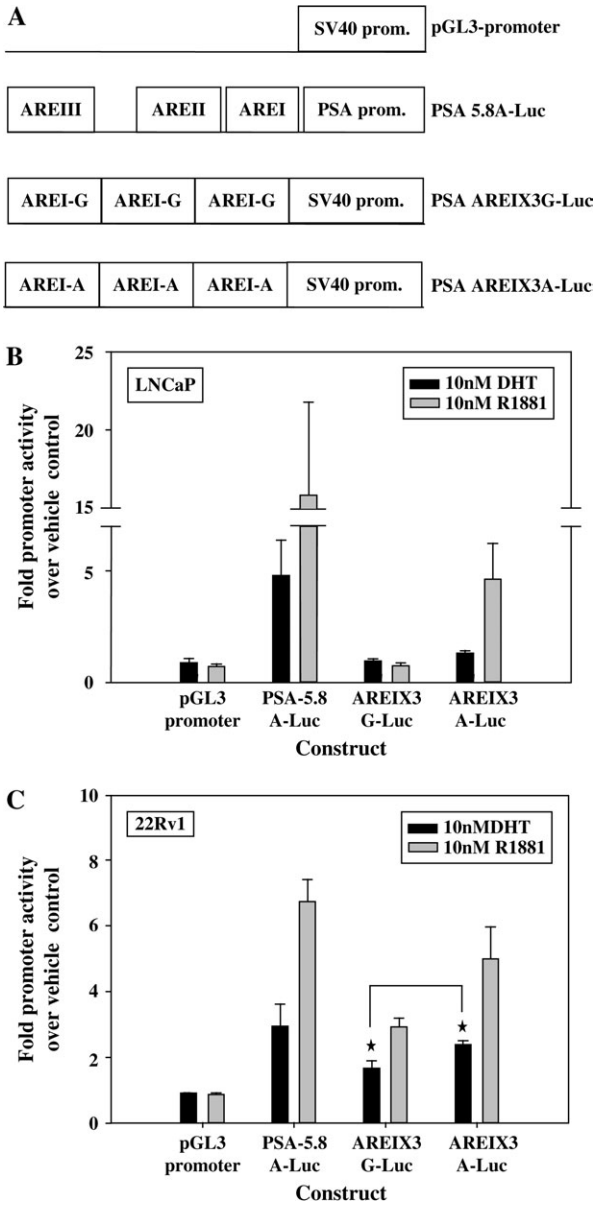

Fig. 3. Luciferase reporter assays assessing differences in androgen responsiveness conferred by the *PSA*G-158A SNP. (**A**) Schematic of constructs used in the analyses. Androgen-regulated promoter assays were carried out in (**B**) LNCaP and (**C**) 22Rv1 cells. Data were normalized to Renilla luciferase activity and expressed as the mean fold increase over vehicle control treatments (1% ethanol) ± SE from three independent experiments. Androgen induction was carried out with both DHT (black bars) and R1881 (grey bars). The asterisks above the bars indicate a statistically significant difference in promoter activity between the two allelic constructs (P = 0.05).

[OR_{AG} = 2.42 (95% CIs = 1.24–4.72); OR_{AA} = 3.07 (95% CIs = 1.41–6.68) and $OR_{AA/AG}$ = 2.62 (95% CIs = 1.39–4.92), P = 0.003].

Discussion

Given the importance of *PSA* and androgens in early prostate (patho)-physiology, and the suggestion that the G-158A polymorphism in AREI of the *PSA* gene might moderate the role of PSA in prostate cancer and be associated with serum levels, we sought to determine the functional significance of this polymorphism using a number of *in vitro* assays. Further to these findings and following conflicting

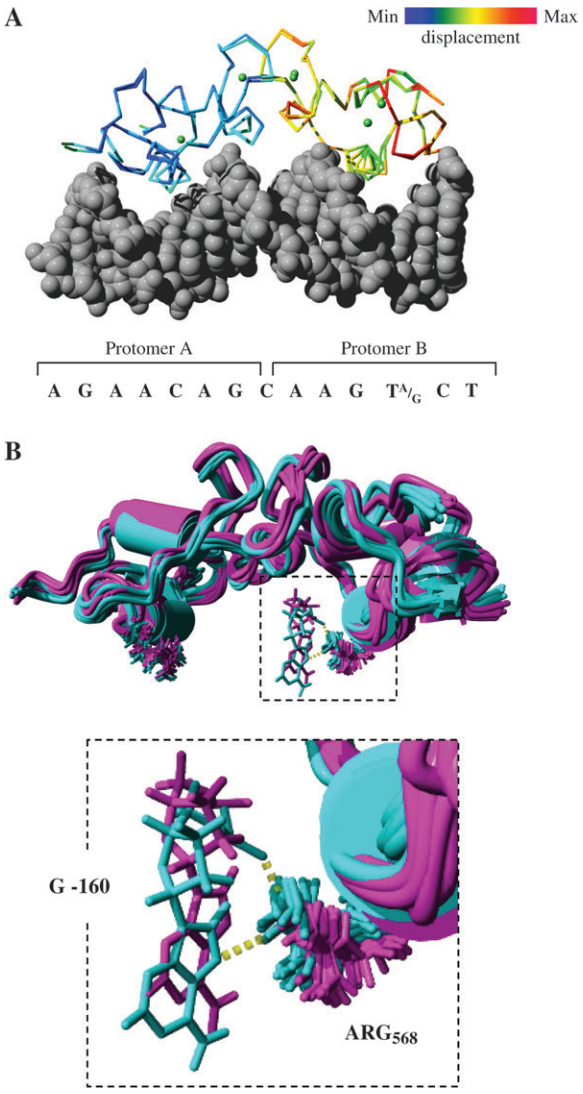

Fig. 4. In *silico* modelling for structural differences conferred by the G-158A SNP. (**A**) Model of AR–DBD protomers binding to −158A and −158G oligonucleotides. The dimeric protein components of both complexes have been superimposed on to a single DNA molecule and are represented by an alpha carbon trace coloured according to displacement from the starting template structure (Structure file PDBid 1I4R) with red indicating maximum displacement and blue minimum displacement. Note that only the B protomer shows significant displacement from the original structure. DNA is represented in space-filling mode and coloured grey, and zinc atoms are coloured green. DNA sequence bound by protomers A and B is also shown schematically. (**B**) Molecular dynamic simulation of the AR–DBD in solution. Molecular dynamics of the −158A and −158G AR–DBD models was followed over a period of 500 ps taking snapshots every 50 ps. Both sets of snapshots have been overlaid using the bound oligonucleotide as point of reference. For the sake of clarity, only G-160 of the bound oligonucleotide is shown. The −158A model snapshots are coloured cyan and the −158G model snapshots are coloured magenta. Greatest deviation between the two sets of structures was observed for ARG_{568} in protomer B. The inset shows a magnified view of this side chain highlighting the two extra hydrogen bonds (yellow dashes) formed between the oligonucleotide G-160 and the B protomer. Note that no significant deflection is observed for this residue in protomer A.

results obtained from previous association studies of the G-158A polymorphism with prostate cancer risk, we conducted a case–control study to examine the association between this polymorphism and prostate cancer risk in Australian men.

Table I. Socio-demographic and clinical characteristics of the study populations

Characteristics[a]	Men with prostate cancer (n = 209) n (%)	Healthy controls (n = 223)[b] n (%)
Age in years (median, range)	68 (50–91)	50 (18–75)
Marital status		
Never married	8 (4)	41 (18)
Married/de facto	179 (86)	151 (68)
Divorced/separated/widowed	18 (9)	20 (9)
Unknown	4 (1)	11 (5)
Family history of prostate cancer		
No	153 (73)	207 (93)
Yes	56 (27)	16 (7)
Vasectomy status		
No	151 (73)	134 (60)
Yes	58 (27)	89 (40)
Smoking status		
Never smoked	94 (45)	121 (54)
Former smoker[c] <20 pack years	43 (20.5)	49 (22)
Former smoker[c] >20 pack years	63 (30)	26 (12)
Current smoker <20 pack years	1 (0.5)	8 (4)
Current smoker >20 pack years	4 (2)	9 (4)
Unknown	4 (2)	10 (4)
Alcohol consumption		
Non-drinker	73 (35)	39 (18)
Drinker	132 (63)	175 (79)
Unknown	4 (2)	9 (3)
Highest education level achieved		
No formal education/ primary school	31 (15)	6 (3)
Secondary school	116 (56)	71 (32)
Technical college	22 (16)	55 (25)
University	33 (10)	81 (36)
Unknown	7 (3)	10 (4)
TNM staging		
pT1	94 (45)	Not applicable
pT2	54 (26)	Not applicable
pT3	32 (15)	Not applicable
pT4	2 (1)	Not applicable
Unknown	27 (13)	Not applicable
Serum PSA at diagnosis		
<4 ng/ml	26 (12)	Not measured
4–10 ng/ml	95 (46)	Not measured
>10 ng/ml	71 (34)	Not measured
Unknown	17 (8)	Not measured
Gleason score (Gleason grade 1 + Gleason grade 2)		
<7	178 (85)	Not applicable
≥7	11 (5)	Not applicable
Unknown	20 (10)	Not applicable

[a]Study characteristics differed significantly between cases and controls ($P < 0.01$).
[b]Approximately half the controls (n = 115) were aged 50 years and over.
[c]Former smokers included participants who reported having ceased smoking at least 1 year before taking part in this study.

In vitro EMSA analysis with purified AR–DBD resulted in a single shift that is consistent with observations by other groups (5,6) and demonstrated that the *PSA* AREI A allele binds with 2-fold greater affinity to the AR–DBD compared with the *PSA* AREI G allele. Our EMSA data are also supported by a previous PCR-based sequence selection study that showed an ∼2-fold greater selection for adenines over guanines at the +5 position (for which the G-158A SNP is positioned) (28).

We then carried out limited proteolysis experiments to determine whether the *PSA* G-158A SNP has an allosteric effect on the AR that may result in altered receptor dimerization/conformation and, in turn, mediate differential cofactor recruitment or affect AREI–receptor complex stability. Trypsin digestion in the 22Rv1 experiments demonstrated an increase in the lowest shift intensity, which coincides with a decrease in intensity for the higher molecular weight shifts

Table II. Genotypic frequencies for the PSA G-158A SNP, with crude and adjusted ORs for prostate cancer risk

Genotype	Genotypic frequencies[a]		Crude OR	95% CIs	Significance *P*	Adjusted OR[b]	95% CIs	Significance *P*
	Healthy controls ($n = 223$)	Men with prostate cancer ($n = 209$)						
GG	0.32	0.17	Referent	—	—	Referent	—	—
AG	0.48	0.53	2.03	1.25–3.28	0.001[c]	2.44	1.23–4.81	0.012[c]
AA	0.20	0.30	2.76	1.59–4.81	—	3.00	1.36–6.52	—
AA/AG[d]	0.68	0.83	2.25	1.42–3.54	0.001	2.61	1.37–4.96	0.004

[a]Both populations were in Hardy–Weinberg equilibrium ($P \geq 0.31$).
[b]Multivariate model adjusted for age (there was no confounding from other variables).
[c]Test for significance of trend.
[d]On a dominant model.

and suggests that the multiple bands in the limited proteolysis assays are intermediate AR co-regulatory complexes. Limited proteolysis experiments with both 22Rv1 and LNCaP nuclear extracts demonstrated that *PSA* AREI-A bound with greater affinity to the AR and its associated co-regulators compared with *PSA* AREI-G. This is consistent with the EMSA analysis that showed a greater binding affinity of the A allele to the AR–DBD. Differences in shift intensities between the two *PSA* AREI alleles after trypsin digestion for some of the shifts in the LNCaP experiments suggest that DNA–receptor stability is also altered through allosteric interactions conferred by the AREI alleles onto the AR, as demonstrated by the higher tolerance of the A allele to trypsin digestion. Hence, altered AREI–AR interactions are probably caused by a combination of differential receptor–AREI-binding affinities and subsequent allosteric effects imparted onto the AR and its co-regulators that ultimately affect receptor–AREI complex stability. Indeed, DNA-mediated allosteric changes in the AR have been previously reported in a limited proteolysis experiment comparing cognate and non-cognate AREs (29). Also of note is that only two common receptor–AREI complexes were detected when comparing bands between the two cell lines. The availability of AR co-regulators might, in part, account for these differences; however, these cell-specific differences do not seem to affect the allosteric effects conferred by the G-158A SNP, suggesting that the A allele may have a more critical role in recruiting and stabilizing the AR and its co-regulators to the *PSA* promoter.

Since there was evidence of altered binding of the *PSA* AREI alleles with the AR and differences in receptor–AREIA/G complex stability, we undertook androgen-regulated reporter assays to determine whether these data translated to a transcriptional response, which is ultimately the end point for AR-targeted genomic effects. The luciferase data demonstrated that the *PSA* AREI A allele was significantly more responsive to both the natural (DHT) and synthetic (R1881) androgens in both LNCaP and 22Rv1 cells. However, it is difficult at this stage to determine whether the G-158A polymorphism has any transcriptional effect in a native context and when acting synergistically with the other *PSA* AREs. Indeed, it was recently reported (15) that variant AREI promoter constructs (which includes AREs I, II and III) were not differentially responsive to androgens. Interestingly, in the same study, the authors also found no functional difference in transcriptional activity for the two AREI alleles using similar tandem constructs as those used in this study. The discrepancy from the data observed in this and the previous functional study are probably caused by several experimental differences between the two studies. For example, in this study, transcriptional activity was normalized to luminescence from Renilla activity, whereas the Rao *et al.* (15) study normalized transcriptional activity to β-galactosidase expression. In our study, cells were starved of steroids for 54 h (48 h serum-starved and 6 h transfection time) as opposed to 24 h for the Rao *et al.* study (15). Hence, incomplete removal of endogenous androgens in the latter study may also account for the observed differences between the two studies. Finally, it is possible that transient expression of the AR in PC-3 cells may be at saturated levels that mask any subtle differences between the two AREI alleles in response to androgens.

Molecular modelling of the *PSA* AREI alleles with an *in silico* representation of the AR–DBD predicted the formation of an extra two hydrogen bonds with the AREI A allele, which would account for the increased binding observed in the EMSA experiments. Although modelling was performed with the DBD in isolation, rather than with an intact AR, the validity of this model was strengthened by correlation of binding energies calculated for the *in silico* representation with those obtained from EMSA experiments. Modelling also suggested that this difference derives from two extra two hydrogen bonds formed between the oligonucleotide and ARG_{568} of protomer B in the dimeric DBD. Interestingly, the extra bonds are not formed with the variant nucleotide at −158A/G, but with a nucleotide further towards the 5′ end of the oligonucleotide (−160G), common to both alleles. This suggests that the sequence of the bound DNA can alter conformation of the bound AR–DBD, which in turn can cause displacements in the bound DNA discrete from variant bases.

Given our observations of the functional importance of the G-158A polymorphism in *PSA* AREI, we were encouraged to look for an association between this polymorphism and prostate cancer risk in an Australian Caucasian population. Using a dominant model, we showed that men with at least one A allele had a 2.6-fold increased risk of developing prostate cancer (95% CIs = 1.37–4.96, $P = 0.004$) and, under a co-dominant model, the AA genotype was associated with a 3-fold increased risk (95% CIs = 1.36–6.52). Although the age distributions of the case and control groups were only crudely matched, this does not explain the basis of the observed association, as subgroup analyses limited to controls aged 50 years and over yielded equivalent age-adjusted results.

The role of the G-158A polymorphism in disease susceptibility has still to be elucidated; however, it has been suggested that it is important only in combination with other gene variants, either within the *PSA* gene itself or in other genes, such as the AR gene (13,16,19,30). Previous studies considering the association of the *PSA* G-158A polymorphism with prostate cancer have been contradictory; however, many of these studies have been conducted in small to moderately sized groups, in populations of different ethnicities, and in studies with varying numbers of cases and controls. An initial, relatively small study (57 cases, 156 controls) found a 3-fold increased risk for advanced prostate cancer in non-Hispanic white men with the GG genotype, and in that population, the risk of developing prostate cancer increased 5-fold when associated with short AR alleles (16). Subsequently, smaller population-based case–control studies by Binnie *et al.* (19) [100 cases, 79 benign prostatic hyperplasia (BPH) controls, 67 population controls] and Chiang *et al.* (20) (122 cases, 84 BPH controls) have confirmed the association between prostate cancer susceptibility and the G allele in Caucasian (19) and Asian (20) populations, as did a larger, sibling-based case–control study (439 cases, 479 controls) in a predominantly white American population (17). Other moderately sized studies have refuted these findings, and showed an association between the A allele

and increased risk of prostate cancer in a Portuguese population (151 case, 127 controls) (14) and an Austrian population (190 cases, 190 BPH controls) (18). However, larger studies have shown no significant association between the G-158A SNP and risk of prostate cancer in a Finnish population (968 cases, 923 population controls) (21) or in a predominantly white American population (591 cases, 538 population controls) (22) or a Japanese population (300 cases, 216 BPH controls, 266 controls) (23).

Since G-158A allele frequencies have been shown to be significantly different between various populations and different studies, the importance of the distribution of the G-158A polymorphism between different ethnic groups, and with respect to prostate cancer susceptibility in these groups, remains a subject of some debate. It has been reported that the alleles occur with approximately equal frequencies (50% A, 50% G) in both white American and African-American populations (15); however, other reports have shown that the G allele is slightly more common among African-American populations (55%), than white American populations (16). Our study revealed that the G allele occurred less frequently (32%) in an Australian white population than in other white populations. It has also been shown that the GG genotype is more common in Japanese-Americans (64%), compared with Hispanic (37%), non-Hispanic white (29%) and black American (24%) populations (13), European white populations (30%) (14) and this Australian white population (32%). Most interestingly, however, the −158A allele has been consistently reported to be underrepresented in Asian populations (13,16,20,23,27), which have a lower incidence of prostate cancer, when compared with Caucasians and African-Americans. The role of other genetic variants in modifying the association with G-158A cannot be discounted, for example, varying numbers of CAG repeats in the AR gene between different ethnic groups may account for discordance between different studies.

It has been proposed that the G-158A polymorphism is associated with circulating serum levels of PSA and/or disease stage (as measured by tumour histology or grade) (11,13,17–20). Higher serum PSA levels have been shown to be associated with the A allele or AA genotype (13,14) in some studies, and with the A/G or G/G genotype in others (11,12). However, we found no association with any of the clinical parameters measured in men with prostate cancer cases, including serum PSA at diagnosis, TNM stage or Gleason score. This may be a function of the relatively small numbers within each clinical subgroup in this study.

In conclusion, we have used a number of sensitive *in vitro* assays to identify a possible functional role for the G-158A polymorphism in AREI of the *PSA* gene, by showing that the A allele confers greater androgen responsiveness and shows a greater affinity for the AR than the G allele. We have also shown an association between the *PSA* AREI-G-158A polymorphism and risk of developing prostate cancer in a well-characterized, albeit moderately sized, Australian Caucasian population. Our results suggest that the G-158A polymorphism may be one mechanism by which *PSA* expression is altered and predisposes to prostate cancer.

Acknowledgements

We thank all the men who participated in this study, as well as Mrs Judi Stewart, Dr Frank Gardiner and Dr David Nicol, Ms Beth Morrison, Mrs Betty Scells, Ms Rachelle Morris, and the staff at the Australian Red Cross Blood Services for their assistance with the collection of clinical information and blood samples, and Mr Andrew Ramsay, Mr Mitchell Lawrence, Dr Joanne Young, Dr Melissa Barker, Ms Melanie Higgins and all the other members of the Molecular Cancer Epidemiology Laboratory for their assistance with processing of blood samples. We also acknowledge the contribution of Prof. Beth Newman in establishing the case–control study resource and thank her for the comments on the manuscript. This study was supported by grants from Queensland University of Technology and the Cooperative Research Centre for Diagnostics. J. A. Clements is a National Health and Medical Research Council's Principal Research Fellow.

Conflict of Interest Statement: None declared.

References

1. Clements,J.A. (1989) The glandular kallikrein family of enzymes: tissue-specific expression and hormonal regulation. *Endocr. Rev.*, **10**, 393–419.
2. Yousef,G.M. *et al*. (2000) Genomic organization of the human kallikrein gene family on chromosome 19q13.3-q13.4. *Biochem. Biophys. Res. Commun.*, **276**, 125–133.
3. Sinha,A.A. *et al*. (1987) Immunoelectron microscopic localization of prostatic-specific antigen in human prostate by the protein A-gold complex. *Cancer*, **60**, 1288–1293.
4. Nelson,C.C. *et al*. (1999) Determinants of DNA sequence specificity of the androgen, progesterone, and glucocorticoid receptors: evidence for differential steroid receptor response elements. *Mol. Endocrinol*, **13**, 2090–2107.
5. Cleutjens,K.B. *et al*. (1997) An androgen response element in a far upstream enhancer region is essential for high, androgen-regulated activity of the prostate-specific antigen promoter. *Mol. Endocrinol*, **11**, 148–161.
6. Cleutjens,K.B. *et al*. (1996) Two androgen response regions cooperate in steroid hormone regulated activity of the prostate-specific antigen promoter. *J. Biol. Chem.*, **271**, 6379–6388.
7. Riegman,P.H. *et al*. (1991) The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. *Mol. Endocrinol*, **5**, 1921–1930.
8. Cramer,S.D. *et al*. (2003) Association between genetic polymorphisms in the prostate-specific antigen gene promoter and serum prostate-specific antigen levels. *J. Natl Cancer Inst.*, **95**, 1044–1053.
9. Thompson,I. *et al*. (2003) Prostate cancer and prostate-specific antigen: the more we know, the less we understand. *J. Natl Cancer Inst.*, **95**, 1027–1028.
10. Rao,A. *et al*. (1999) Identification of a polymorphism in the ARE I region of the PSA promoter. *Proc. Am. Assoc. Cancer Res.*, **40**, 60.
11. Schatzl,G. *et al*. (2005) Polymorphism in ARE-I region of prostate-specific antigen gene associated with low serum testosterone level and high-grade prostate cancer. *Urology*, **65**, 1141–1145.
12. Xu,J. *et al*. (2002) Association studies of serum prostate-specific antigen levels and the genetic polymorphisms at the androgen receptor and prostate-specific antigen genes. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 664–669.
13. Xue,W.M. *et al*. (2001) Genetic determinants of serum prostate-specific antigen levels in healthy men from a multiethnic cohort. *Cancer Epidemiol. Biomarkers Prev.*, **10**, 575–579.
14. Medeiros,R. *et al*. (2002) Linkage between polymorphisms in the prostate specific antigen ARE1 gene region, prostate cancer risk, and circulating tumor cells. *Prostate*, **53**, 88–94.
15. Rao,A. *et al*. (2003) Analysis of G/A polymorphism in the androgen response element I of the PSA gene and its interactions with the androgen receptor polymorphisms. *Urology*, **61**, 864–869.
16. Xue,W. *et al*. (2000) Susceptibility to prostate cancer: interaction between genotypes at the androgen receptor and prostate-specific antigen loci. *Cancer Res.*, **60**, 839–841.
17. Cicek,M.S. *et al*. (2005) Role of androgen metabolism genes CYP1B1, PSA/KLK3, and CYP11alpha in prostate cancer risk and aggressiveness. *Cancer Epidemiol. Biomarkers Prev.*, **14**, 2173–2177.
18. Gsur,A. *et al*. (2002) Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis*, **23**, 1647–1651.
19. Binnie,M.C. *et al*. (2005) Polymorphic forms of prostate specific antigen and their interaction with androgen receptor trinucleotide repeats in prostate cancer. *Prostate*, **63**, 309–315.
20. Chiang,C.H. *et al*. (2004) The impact of polymorphism on prostate specific antigen gene on the risk, tumor volume and pathological stage of prostate cancer. *J. Urol.*, **171**, 1529–1532.
21. Mononen,N. *et al*. (2006) Profiling genetic variation along the androgen biosynthesis and metabolism pathways implicates several single nucleotide polymorphisms and their combinations as prostate cancer risk factors. *Cancer Res.*, **66**, 743–747.
22. Salinas,C.A. *et al*. (2005) Polymorphisms in the androgen receptor and the prostate-specific antigen genes and prostate cancer risk. *Prostate*, **65**, 58–65.
23. Wang,L.Z. *et al*. (2003) Polymorphisms in prostate-specific antigen (PSA) gene, risk of prostate cancer, and serum PSA levels in Japanese population. *Cancer Lett.*, **202**, 53–59.
24. Krieger,E. *et al*. (2004) Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins*, **57**, 678–683.
25. Shaffer,P.L. *et al*. (2004) Structural basis of androgen receptor binding to selective androgen response elements. *Proc. Natl Acad. Sci. USA*, **101**, 4758–4763.

26. Miller,S.A. *et al*. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, **16**, 1215.
27. Liu,J. *et al*. (2003) Polymorphisms of prostate-specific antigen gene promoter: determination from cord blood collected on filter paper. *Ann. Clin. Lab. Sci.*, **33**, 429–434.
28. Roche,P.J. *et al*. (1992) A consensus DNA-binding site for the androgen receptor. *Mol. Endocrinol*, **6**, 2229–2235.
29. Geserick,C. *et al*. (2003) Differential modulation of androgen receptor action by deoxyribonucleic acid response elements. *Mol. Endocrinol*, **17**, 1738–1750.
30. Sieh,W. *et al*. (2006) Genetic susceptibility to prostate cancer: prostate-specific antigen and its interaction with the androgen receptor (United States). *Cancer Causes Control*, **17**, 187–97.

Received August 14, 2006; revised October 27, 2006; accepted November 24, 2006