ORIGINAL ARTICLE

The androgen receptor CAG and GGN repeat polymorphisms and prostate cancer susceptibility in African-American men: results from the Flint Men's Health Study

Ethan M. Lange · Aruna V. Sarma · Anna Ray ·
Yunfei Wang · Lindsey A. Ho · Sarah A. Anderson ·
Julie M. Cunningham · Kathleen A. Cooney

Abstract Repeat lengths of the CAG and GGN microsatellites in exon 1 of the androgen receptor (*AR*) gene have been hypothesized to be associated with prostate cancer risk. In vitro studies have showed an inverse association between *AR* CAG and GGN repeat length and activity levels of the AR product. It is known that men of African descent have a higher incidence of and greater mortality from prostate cancer than men of Caucasian or Asian descent and, on average, a smaller number of repeats at *AR* CAG and GGN. Consistent with these findings, studies have also found increased AR protein expression levels in benign prostatic hyperplasia and prostatic diseased tissues from men of African descent. Despite these findings, limited studies have been conducted to evaluate the association between repeat lengths at *AR* CAG and prostate cancer risk in African Americans. Our study is the first such study to examine whether repeat length of the *AR* GGN repeat is associated with prostate cancer risk in African Americans. We found no evidence for an association between *AR* CAG or GGN repeat lengths and prostate cancer risk in a population-based sample of African Americans.

Keywords Prostate cancer · Androgen receptor · CAG repeat · GGN repeat · Association

E. M. Lange (✉) · Y. Wang
Department of Genetics, University of North Carolina-Chapel Hill, 4300D MBRB, CB# 7264, 111 Mason Farm Road, Chapel Hill, NC 27599-7264, USA
e-mail: elange@med.unc.edu

E. M. Lange · L. A. Ho
Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA

E. M. Lange
Carolina Center for Genome Sciences,
University of North Carolina, Chapel Hill, NC, USA

A. V. Sarma
Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA

A. Ray · K. A. Cooney
Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA

S. A. Anderson · J. M. Cunningham
Department of Laboratory Medicine and Pathology, Mayo Clinic/Foundation, Rochester, MN, USA

K. A. Cooney
Department of Urology, University of Michigan, Ann Arbor, MI, USA

K. A. Cooney
University of Michigan Comprehensive Cancer Center, Ann Arbor, MI, USA

Introduction

In 2006, prostate cancer was the most commonly diagnosed non-skin cancer among men in the United States, with an estimated 218,890 new cases, and a leading cause of cancer-related mortality, with an estimated 27,050 related deaths (Jemal et al. 2007). Increasing age, positive family history and African ancestry are known risk factors for prostate cancer (Bostwick et al. 2004).

The clustering of prostate cancer cases among families has motivated the search for genetic risk factors for the disease. One of the most studied genes implicated in prostate cancer susceptibility is the androgen receptor (*AR*) gene on chromosome Xq11-12. The *AR* gene product regulates expression of the genes necessary for growth and

development of many target tissues, including male reproductive organs. Variant forms of two different microsatellite polymorphisms [the polyglutamine $(CAG)_n$ repeat and the imperfect polyglycine $(GGN)_n$ repeat, both located in exon 1 of *AR*] have been shown to alter the biological function of the *AR* gene and consequently have been hypothesized to modify the risk for developing prostate cancer. Specifically, in vitro studies (Beilin et al. 2000; Chamberlain et al. 1994; Kazemi-Esfarjani et al. 1995; Tut et al. 1997; Ding et al. 2004; Ding et al. 2005) have demonstrated an inverse relationship between the length of both repeats and AR activity levels. The recent article by Rajender et al. (2007) provides a nice review on the *AR* CAG and GGN repeats with respect to their function and their statistical association with a wide range of clinical phenotypes. Studies comparing the distribution of allele sizes for the CAG and GGN microsatellite repeat polymorphisms between different populations have noted shorter repeat lengths, on average, in men of African versus Caucasian descent (Kittles et al. 2001; Esteban et al. 2006). Consistent with the in vitro studies showing increased activity of the AR product with shorter *AR* CAG and GGN repeat alleles, studies have also found increased AR protein expression levels in prostatic tissues from men of African descent (Gaston et al. 2003; Olapade-Olaopa et al. 2004). Specifically, AR protein expression levels were estimated to be 22% higher in benign prostate cancer tissues and 81% higher in malignant prostate cancer tissues in African Americans versus Caucasian Americans (Gaston et al. 2003).

Through a literature search, we have identified over 30 studies that have evaluated the association between the *AR* CAG and GGN repeats and prostate cancer. Results from many of these studies have been summarized in a meta-analysis (Zeegers et al. 2004). These studies have focused primarily on men of Caucasian descent. Typical of genetic association studies for complex traits, the results from these studies have varied considerably. Taken together, the cumulative results across these different studies suggest that if there is an effect of short alleles at the *AR* CAG and GGN repeats on prostate cancer risk, then the magnitude of the differential risk, at least in Caucasian men, is likely small.

African-American men have an approximately 1.6-fold greater chance of being diagnosed with prostate cancer compared to Caucasian men and a 2.4-fold greater chance of dying from the disease (Jemal et al. 2007). In addition, African-American men are more frequently diagnosed with higher tumor volume, more advanced tumor stage, higher Gleason grade and higher prostate-specific antigen levels (Brawn et al. 1993; Vijayakumar et al. 1998; Fowler and Bigler 1999; Moul et al. 1999; Powell et al. 1999; Fowler et al. 2000; Thompson et al. 2001). These findings have

suggested that prostate cancer in African-American men may involve different etiological factors and may be more biologically aggressive. Given the strong genetic heterogeneity of prostate cancer, different levels of background risk factors and plausibly unique genetic and environmental interactions, it is important to study the effects of *AR* CAG and GGN repeat lengths on prostate cancer susceptibility directly in African-American men rather than relying on results from extensive studies in Caucasian men.

Unfortunately, despite their increased risk for developing the disease, very limited studies regarding the risk of these *AR* repeat polymorphisms on prostate cancer in African Americans have been conducted. A small study on 20 men of African descent diagnosed with prostate cancer and 20 healthy controls found no evidence for an association between *AR* CAG allele size and prostate cancer risk (Panz et al. 2001). Similarly, a study on 118 African-American men diagnosed with prostate cancer and 567 African-American controls revealed no evidence for an association between *AR* CAG repeat length and prostate cancer (Gilligan et al. 2004). A multiethnic cohort study with 635 African-American prostate patients and 664 African-American controls also failed to identify any association between *AR* CAG repeat length and prostate cancer (Freedman et al. 2005).

Herein, we evaluate the association between the *AR* CAG and *AR* GGN repeat polymorphisms and prostate cancer in a community-based sample of 471 African-American men from Flint, Michigan (Cooney et al. 2001). To our knowledge, this is the first study to assess the risk of the *AR* GGN repeat on prostate cancer in African-American men. We found no evidence to support the association between either the *AR* CAG or *AR* GGN repeat polymorphisms and prostate cancer in this population.

Materials and methods

Subjects

African-American subjects from this study were part of the Flint Men's Health Study (FMHS) (Cooney et al, 2001). The FMHS is a community-based study of prostate cancer in African-American men between the ages of 40 and 79 years. In 1996, 730 men were recruited to participate in the study from a probability sample residing in the city of Flint and surrounding communities in Genesee County, Michigan. Subjects completed a detailed in-home interview that collected information on socio-demographics, potential risk factors for prostate cancer and a complete medical history. Subjects were also asked to participate in a clinical examination that included measurement of serum PSA (free and total) and a comprehensive urologic examination.

Men with an elevated total PSA (>4.0 ng/ml) or an abnormal digital rectal exam were referred for prostate biopsy. Of the 730 men who completed the initial interview, 379 participated in the clinical exam. A total of ten subjects were diagnosed with prostate cancer as a consequence of the protocol, which resulted in a final control sample of 369 men. Attempts were made to follow the study participants, and in the 5 years after control recruitment, an additional 18 control men were diagnosed with prostate cancer. For this study, a sufficient DNA sample was available for genotyping on 342 controls.

Cases were recruited from the same community from 1999 to July 2002. Men who were between the ages of 40 and 79 years at the time of prostate cancer diagnosis (between 1995 and 2002) were eligible to participate in the study. Patients also completed a detailed epidemiologic interview and provided a blood sample. Medical records were reviewed to extract information related to prostate cancer diagnosis including clinical and pathologic stage, Gleason grade, prediagnostic PSA and treatment. A total of 136 patients were ultimately recruited to participate in the study (including the control men who were diagnosed with prostate cancer through participation in the study, $n = 10$). For this study, sufficient DNA samples were available for genotyping on 131 cases. Informed consent was obtained from all study participants, and the research protocol was approved by the Institutional Review Board of the University of Michigan.

Genotyping

For both cases and controls, genomic DNA was isolated from whole blood using the Puregene DNA Purification Kit (Gentra Systems, Minneapolis, MN). The number of *AR* CAG and GGN DNA repeats was determined by PCR-based fragment analysis, using fluorescently labeled primers (Roberts et al. 2004). Briefly, each 15 μl reaction contained 15 ng genomic DNA, 2 mM MgCl$_2$, 200 μM dNTPs, 0.67 μM each primers and 0.5 U Amplitaq Gold (Applied Biosystems, Foster City, CA). The annealing temperatures were 55°C for *AR* CAG repeats and 58°C for *AR* GGN repeats. PCR products were resolved on an ABI 3100 DNA sequencer (Applied Biosystems), each capillary is calibrated using internal reference standards and control samples were included in each plate to ensure accuracy of genotypes.

Genotype was scored successfully for 130/131 and 129/131 cases for *AR* CAG and GGN, respectively. Genotype was scored successfully for both *AR* CAG and GGN for 340/342 controls; 2 controls did not successfully genotype at either *AR* CAG or GGN. In total, 131 patients and 340 controls had genotype data available on at least one of the two *AR* repeats.

Statistical methods

To measure the strength of dependence (of allele size) and linkage disequilibrium between the CAG and GGN repeats, we calculated Spearman's correlation coefficient and used Lewontin's D' (Lewontin 1964) modified for multiple alleles (Hedrick 1987) separately for both the case and control samples. For two-allele markers, D' is the standardized disequilibrium value that takes the usual disequilibrium coefficient $P(A_iB_j) - P(A_i)P(B_j)$ and divides it by its maximal possible value. Given multiple alleles, we calculate the weighted average of the D' values where the weights are the products of the corresponding allele frequencies. That is,

$$D' = \sum_i \sum_j p_i q_j \left| D'_{ij} \right|,$$

where p_i and q_j are allele frequencies at the two loci of interest, and D_{ij}' is the standardized disequilibrium coefficient based on alleles A_i and B_j. Statistical significance of the magnitude of the estimated D' values was assessed using a permutation test in which, under the null distribution (i.e., linkage equilibrium), within a sample the alleles at the two repeats were randomly shuffled between individuals independently.

To evaluate the haplotype diversity in this African-American sample, we constructed the observed haplotype frequency distribution in the combined sample of 468 men with genotype data available at both *AR* CAG and GGN. We determined the median allele size for each repeat in the combined sample and used these observed medians as allele size thresholds to partition the haplotypes from the complete sample into four groups based on allele size combinations at the two repeats. We then performed an additional test of independence of allele sizes at the two repeats by calculating the expected number of haplotypes for each of the four groups and used a Pearson's chi-square test to evaluate whether the observed number of haplotypes in each group was consistent with the expected numbers.

We used unconditional multivariable logistic regression models to assess the association between *AR* CAG and GGN repeat lengths and prostate cancer. Two levels of covariate adjustment were made to all models: (1) age only and (2) age and estimated proportion of African ancestry. Approximately half of the FMHS control men were tested at multiple time points for prostate cancer. To avoid lead-time bias in the multivariable analyses, age was calculated based on the same date for all cases and controls. This date was the most recent follow-up date from the entire sample, with the exception that age at death was used for the 37 controls that died prior to this date. Estimated proportion of African ancestry for each study participant had been obtained previously (Amundadottir et al. 2006) using the

statistical software Structure (Pritchard 2000). *AR* CAG and GGN repeat length were analyzed as both continuous measures or dichotomized based on repeat length thresholds previously suggested in the *AR*-prostate cancer literature (two cut-off values were considered for CAG: ≤21 vs. >21 and ≤22 vs. >22 and one cut-off value for GGN: ≤16 vs. >16). We analyzed the effects of CAG and GGN repeats separately and jointly. In addition, we tested for interaction effects between the *AR* CAG and GGN repeats on prostate cancer risk. Finally, *t* tests were used to assess statistical significance of observed differences in means for age and estimated proportion of African ancestry between cases and controls. All analyses were performed using the SAS, version 9.1.3, statistical software package (SAS Institute, Cary, NC), and all tests, unless otherwise stated, were evaluated using a two-sided hypothesis test.

Results

The total sample consisted of 471 (131 prostate cancer cases, 340 disease-free controls) African-American subjects. Mean age overall was 63.5 years (standard deviation or SD = 10.0) with patients being older than controls (patient mean age = 67.2 years, SD = 8.6; control mean age = 62.1 years, SD = 10.1; $P < 0.0001$). A family history of prostate cancer in a first-degree relative was reported by 21.4% of the patients and 17.0% of the controls. Based on the definition from the International Consortium for Prostate Cancer Genetics (Schaid et al. 2006), 72.5% of the patients were assessed to have clinically aggressive prostate cancer. Sixty percent of cases had cancers with a Gleason score of 7, and 11% had cancers with a Gleason score ranging from 8 to 10. There was no statistical difference in mean proportion of African descent between cases (mean proportion = 0.705, SD = 0.077) and controls (mean proportion = 0.707, SD = 0.077).

The observed allele frequencies for cases and controls for *AR* CAG and GGN are presented graphically in Figs. 1 and 2, respectively. The average number of repeats was similar for cases and controls for both the *AR* CAG (case mean = 19.92, SD = 3.37, median = 19; control mean = 19.91, SD = 3.47, median = 20) and GGN (case mean = 15.76, SD = 2.10, median = 16; control mean = 15.41, SD = 2.22, median = 16) polymorphisms. *AR* CAG and GGN repeat lengths were significantly negatively correlated in both cases (Spearman's correlation = −0.17, $P = 0.05$) and controls (Spearman's correlation = −0.17, $P = 0.002$), indicating that shorter repeats for *AR* CAG are on haplotypes containing longer GGN repeats, and visa versa, in this population. Consistent with these findings, *AR* CAG and GGN were found to be in significant linkage disequilibrium in both cases (estimated
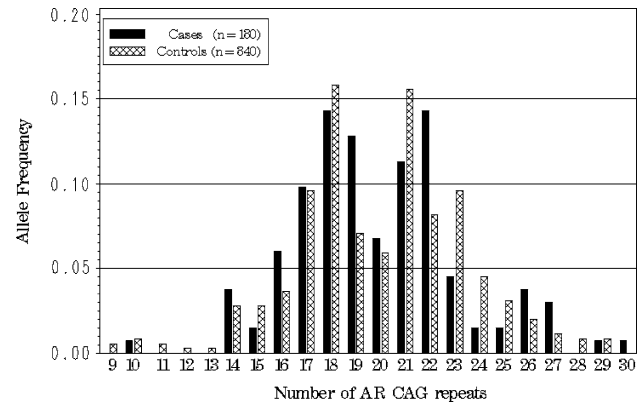

Fig. 1 Allele frequency distribution for AR CAG repeat
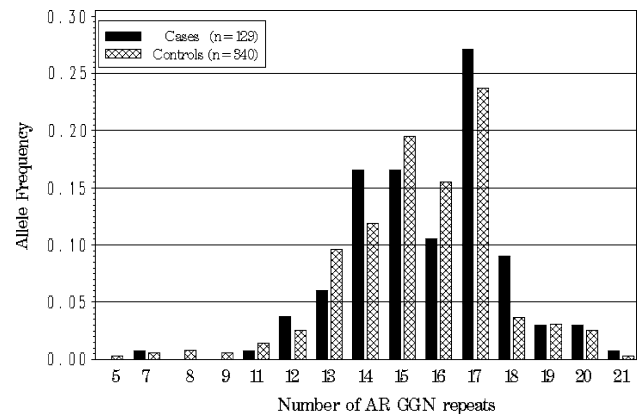

Fig. 2 Allele frequency distribution for AR GGN repeat

$D' = 0.55$, $P < 0.0001$) and controls (estimated $D' = 0.41$, $P < 0.0001$).

A total of 112 distinct haplotypes were observed among our complete sample of 468 men with genotype data on both *AR* CAG and GGN, with the haplotype defined by *AR* CAG = 18 and GGN = 17 the most common observed haplotype (frequency = 0.058). Only 11 different haplotypes were observed more than ten times each, while 41 haplotypes were observed just once. We observed 129 (CAG ≤ 20, GGN ≤ 16), 122 (CAG ≤ 20, GGN > 16), 164 (CAG > 20, GGN ≤ 16) and 53 (CAG > 20, GGN > 16) haplotypes in groupings based on the observed median repeat length values of *AR* CAG (median = 20) and GGN (median = 16). Consistent with the observed negative correlation between allele lengths at *AR* CAG and GGN, these observed counts were significantly different than the expected number of haplotypes in these groupings (157, 94, 136 and 81, respectively) under the null hypothesis of independence ($P < 0.0001$).

Results from the logistic regression models are presented in Table 1. Age was a significant risk factor for prostate cancer ($P < 0.0001$) in all models. Estimated proportion of African descent was not a significant risk

Table 1 Main effects for *AR* CAG and GGN modeled independently

Repeat	Adjustment	Model	OR[a]	95% CI	*P* value
CAG	Age	Continuous[b]	1.00	(0.94, 1.06)	0.86
		≤21 versus >21	0.94	(0.60, 1.47)	0.77
		≤22 versus >22	0.64	(0.37, 1.11)	0.11
CAG	Age + %African descent	Continuous	1.00	(0.94, 1.06)	0.89
		≤21 versus >21	0.93	(0.60, 1.47)	0.77
		≤22 versus >22	0.65	(0.37, 1.11)	0.12
GGN	Age	Continuous	1.08	(0.98, 1.19)	0.11
		≤16 vs. >16	1.42	(0.93, 2.17)	0.11
GGN	Age + %African descent	Continuous	1.08	(0.98, 1.19)	0.11
		≤16 versus >16	1.41	(0.92, 2.16)	0.11

[a] An OR <1 suggests shorter alleles are associated with increased odds of prostate cancer

[b] OR for genotype analyzed as a continuous variable is presented for an increase of one repeat unit

factor in any models and, as demonstrated in Table 1, modified the observed effects of genotype negligibly. No evidence for an association between *AR* CAG repeat lengths and prostate cancer was detected when modeling *AR* CAG repeat length as a continuous variable or as a dichotomous variable with allele cutoff thresholds of ≤21 and ≤22 repeats, after adjustment for age or age and estimated proportion African ancestry. Similarly, no statistically significant association was found between *AR* GGN repeat length and prostate cancer regardless of whether GGN repeat length was treated as a continuous measure or as a dichotomous variable defined by an allele cutoff threshold value of ≤16 repeats. Given the a priori hypothesis that shorter alleles at both *AR* CAG and GGN increase risk for prostate cancer, we get a suggestive one-sided *P* value ($P = 0.055$) when using a cutoff threshold of ≤22 repeats for *AR* CAG. Applying a similar one-sided hypothesis test to *AR* GGN would result in a decreased estimate of statistical significance (versus the two-sided test) given that we observed modestly longer repeat lengths for *AR* GGN among cases. Modeling *AR* CAG (≤22 vs. >22) and GGN (≤16 vs. >16) jointly reduced the estimated significance for CAG modestly ($P = 0.18$). No evidence for a significant interaction between *AR* CAG and GGN was detected ($P = 0.49$). Finally, we found no evidence for an association with prostate cancer when evaluating men with *AR* CAG ≤22 and GGN ≤16 ($P = 0.47$) or men with *AR* CAG ≤22 or GGN ≤16 ($P = 0.42$).

Discussion

Our results, from a population-based sample of 131 African-American men diagnosed with prostate cancer and 340 screened African-American male controls, showed no significant evidence of an association between shorter alleles at *AR* CAG or GGN and increased risk of prostate cancer. In fact, we observed modestly longer GGN repeat lengths among our cases. Our data, combined with three previous reports, suggest that the observation of shorter alleles at *AR* CAG in African Americans does not significantly account for increased prostate cancer risk in African Americans and does not appear to explain the difference in incidence of the disease between men of African and Caucasian descent. Our study is the first study to evaluate the effect of allele length for *AR* GGN on prostate cancer risk in African Americans. Our findings suggest that a shorter number of repeats at *AR* GGN do not have a major effect on prostate cancer susceptibility. Clearly larger studies on African Americans will be necessary to have sufficient power to conclusively evaluate whether there are any mild effects of the *AR* CAG and GGN polymorphisms on prostate cancer risk in this population.

The *AR* CAG and GGN repeats are only 1,176-bp apart (for a CAG repeat length of 22 repeats), suggesting these two microsatellites are likely to be in linkage disequilibrium or LD (Salinas et al. 2005). Kittles et al. (2001) noted increased haplotype diversity for these repeats in individuals of African descent and computed pair-wise estimates of LD for all possible combinations of allele sizes at the two repeats. Their results suggested that the allele lengths at these repeats are not independent in African populations, but that there was no evidence of LD in the other populations considered (though it should be noted that sample sizes were considerably smaller for the other populations). Given the pair-wise analytic strategy, it was difficult to determine whether there was any consistent pattern of shorter alleles at one repeat being associated with longer alleles at the other. Hsing et al. (2000) found no evidence for any correlation between allele sizes at the two repeats in a sample of 190 prostate cancer cases and 304 controls

from China (Spearman's correlation $= -0.03$, $P > 0.05$), and Salinas et al. (2005) reported no evidence for LD in a sample of 455 Caucasian controls ($D' = 0.11$). However, some evidence for a negative correlation between allele sizes at these two repeats has been reported (Irvine et al. 1995; Correa-Cerro et al. 1999; Chang et al. 2002). In the largest of these studies, Chang et al. (2002) found strong evidence for LD ($P = 0.0003$) in a sample of approximately 350 unrelated Caucasian sporadic prostate cancer cases and controls. To summarize, the evaluation of LD between these two *AR* microsatellites has not been performed using uniform methodology across studies, and it is therefore difficult to evaluate the direction and overall significance of LD between these two repeats in African Americans as well as other racial groups. We have performed an extensive analysis of haplotype structure and LD between *AR* CAG and GGN repeats in our African-American sample and found considerable haplotype diversity in this sample and strong evidence for a negative correlation between allele sizes at the two repeats. One implication of these findings is that results, at least in African Americans, from association studies using *AR* CAG and GGN repeat lengths are not independent and that future studies should consider modeling the effects of the two repeats jointly in addition to analyzing their effects individually.

References

Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Cazier JB, Sainz J, Jakobsdottir M, Kostic J, Magnusdottir DN, Ghosh S, Agnarsson K, Birgisdottir B, Le Roux L, Olafsdottir A, Blondal T, Andresdottir M, Gretarsdottir OS, Bergthorsson JT, Gudbjartsson D, Gylfason A, Thorleifsson G, Manolescu A, Kristjansson K, Geirsson G, Isaksson H, Douglas J, Johansson JE, Balter K, Wiklund F, Montie JE, Yu X, Suarez BK, Ober C, Cooney KA, Gronberg H, Catalona WJ, Einarsson GV, Barkardottir RB, Gulcher JR, Kong A, Thorsteinsdottir U, Stefansson K (2006) A common variant associated with prostate cancer in European and African populations. Nat Genet 38:652–658

Beilin J, Ball EM, Favaloro JM, Zajac JD (2000) Effect of the androgen receptor CAG repeat polymorphism on transcriptional activity: specificity in prostate and non-prostate cell lines. J Mol Endocrinol 25:85–96

Bostwick DG, Burke HB, Djakiew D, Euling S, Ho S, Landolph J, Morrison H, Sonawane B, Shifflett T, Waters DJ, Timms B (2004) Human prostate cancer risk factors. Cancer 101:2371–2490

Brawn PN, Johnson EH, Kuhl DL, Riggs MW, Speights VO, Johnson CF 3rd, Pandya PP, Lind ML Bell NF (1993) Stage at presentation and survival of white and black patients with prostate carcinoma. Cancer 71:2569–2573

Chamberlain NL, Driver ED, Miesfeld RL (1994) The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acids Res 22:3181–3186

Chang BL, Zheng SL, Hawkins GA, Isaacs SD, Wiley KE, Turner A, Carpten JD, Bleeker ER, Walsh PC, Trent JM, Meyers DA, Isaacs WB, Xu J (2002) Polymorphic GGC repeats in the androgen receptor gene are associated with hereditary and sporadic prostate cancer risk. Hum Genet 110:122–129

Cooney KA, Strawderman MS, Wojno KJ, Derr KM, Taylor A, Alscer KH, Heeringa SG, Taylor JM, Wei JT, Montie JE, Schottendfeld D (2001) Age-specific distribution of serum prostate-specific antigen in a community-based study of African American men. Urology 57:91–96

Correa-Cerro L, Wohr G, Haussler J, Berthon P, Drelon E, Mangin P, Fournier G, Cussenot O, Kraus P, Just W, Paiss T, Cantu JM, Vogel W (1999) (CAG)$_n$CAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French–German population. Eur J Hum Genet 7:357–362

Ding D, Xu L, Menon M, Reddy GP, Barrack ER (2004) Effect of short CAG (glutamine) repeat on human androgen receptor function. Prostate 58:23–32

Ding D, Xu L, Menon M, Reddy GP, Barrack ER (2005) Effect of GGC (glycine) repeat length polymorphism in the human androgen receptor on androgen function. Prostate 62:133–139

Esteban E, Rodon N, Via M, Gonzalez-Perez E, Santamaria J, Dugoujon JM, Chennawi FE, Melhaoui M, Cherkaoui M, Vona G, Harich N, Moral P (2006) Androgen receptor CAG and GGC polymorphisms in Mediterraneans: repeat dynamics and population relationships. J Hum Genet 51:129–136

Fowler JE Jr, Bigler SA (1999) A prospective study of the serum prostate specific antigen concentrations and Gleason histologic scores of black and white men with prostate cancer. Cancer 86:836–841

Fowler JE Jr, Bigler SA, Bowman G, Kilambi NK (2000) Race and cause specific survival with prostate cancer: influence of clinical stage, Gleason score, age and treatment. J Urol 163:150–151

Freedman ML, Pearce CL, Penney KL, Hirschhorn JN, Kolonel LN, Henderson BE, Altshuler D (2005) Systematic evaluation of genetic variation at the androgen receptor locus and risk of prostate cancer in a multiethnic cohort study. Am J Hum Genet 76:82–90

Gaston KE, Kim D, Singh S, Ford OH, Mohler JL (2003) Racial differences in androgen receptor protein expression in men with clinically localized prostate cancer. J Urol 170:990–993

Gilligan T, Manola J, Sartor O, Weinrich SP, Moul JW, Kantoff PW (2004) Absence of a correlation of androgen receptor gene CAG repeat length and prostate cancer risk in an African-American population. Clin Prostate Cancer 3:98–103

Hedrick PW (1987) Gametic disequilibrium measures: proceed with caution. Genetics 117:331–341

Hsing AW, Gao YT, Wu G, Wang X, Deng J, Chen YL, Sesterhenn IA, Mostofi FK, Benichou J, Chang C (2000) Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case–control study in China. Cancer Res 60:5111–5116

Irvine RA, Yu MC, Ross RK, Coetzee GA (1995) The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. Cancer Res 55:1937–1940

Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ (2007) Cancer Statistics, 2007. CA Cancer J Clin 57(1):43–66

Kazemi-Esfarjani P, Trifiro MA, Pinsky L (1995) Evidence for a repressive function of the long polyglutamine tract in the human

androgen receptor: possible pathogenetic relevance for the (CAG)$_n$-expanded neuronopathies. 4:523–527

Kittles RA, Young D, Weinrich S, Hudson J, Argyropoulos G, Ukoli F, Adams-Campbell L, Dunston GM (2001) Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: implications for prostate cancer risk. Hum Genet 109:253–261

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 49:49–67

Moul JW, Connelly RR, Mooneyhan RM, Zhang W, Sesterhenn IA, Mostofi FK, McLeod DG (1999) Racial differences in tumor volume and prostate specific antigen among radical prostatectomy patients. J Urol 162:394–397

Olapade-Olaopa EO, Muronda CA, MacKay EH, Danso AP, Sandhu DP, Terry TR, Habib FK (2004) Androgen receptor protein expression in prostatic tissues in Black and Caucasian men. Prostate 59:460–468

Panz VR, Joffe BI, Spitz I, Lindenberg T, Farkas A, Haffejee M (2001) Tandem CAG repeats of the androgen receptor gene and prostate cancer risk in black and white men. Endocrine 15:213–216

Powell IJ, Banerjee M, Sakr W, Grignon D, Wood DP Jr, Novallo M, Pontes E (1999) Should African-American men be tested for prostate carcinoma at an earlier age than white men? Cancer 85:472–477

Pritchard JK (2000) Association mapping in structured populations. Am J Hum Genet 67:170–181

Rajender S, Singh L, Thangaraj K (2007) Phenotypic heterogeneity of mutations in androgen receptor gene. Asian J Androl 9:147–179

Roberts RO, Bergstralh EJ, Cunningham JM, Hebbring SJ, Thibodeau SN, Lieber MM, Jacobsen SJ (2004) Androgen receptor gene polymorphisms and increased risk of urologic measures of benign prostatic hyperplasia. Am J Epidemiol 159:269–276

Salinas CA, Austin MA, Ostrander EO, Stanford JL (2005) Polymorphisms in the androgen receptor and the prostate-specific antigen genes and prostate cancer risk. Prostate 65:58–65

Schaid DJ, McDonnell SK, Zarfas KE, Cunningham JM, Hebbring S, Thibodeau SN, Eeles RA, Easton DF, Foulkes WD, Simard J, Giles GG, Hopper JL, Mahle L, Badzioch M, Bishop T, Evans C, Edwards S, Meitz J, Bullock S, Hope Q, Guy M, Hsieh C-L, Halpern J, Balise RR, Oakley-Girvan I, Whittemore AS, Xu J, Dimitrov L, Chang BL, Adams TS, Turner AR, Meyers DA, Friedrichsen DM, Deutsch K, Kolb S, Janer M, Hood L, Ostrander EA, Stanford JL, Ewing CM, Gielzak M, Isaacs SD, Walsh PC, Wiley KE, Isaacs WB, Lange EM, Ho LA, Beebe-Dimmer JL, Wood DP, Cooney KA, Seminara D, Ikonen T, Baffoe-Bonnie A, Fredriksson H, Matikainen MP, Tammela TLJ, Bailey-Wilson J, Schleutker J, Maier C, Herkommer K, Hoegel JJ, Vogel W, Paiss T, Wiklund F, Emanuelsson M, Stenman E, Jonsson BA, Gronberg H, Camp NJ, Farnham J, Cannon-Albright LA, Catalona WJ, Suarez BK, Roehl KA (2006) Pooled genome linkage scan of aggressive prostate cancer: results from the International Consortium for Prostate Cancer Genetics. Hum Genet 120:471–485

Thompson I, Tangen C, Tolcher A, Crawford E, Eisenberger M, Moinour C (2001) Association of African-American ethnic background with survival in men with metastatic prostate cancer. J Natl Cancer Inst 93:219–225

Tut TG, Ghadessy FJ, Trifiro MA, Pinsky L, Yong EL (1997) Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. J Clin Endocrinol Metab 82:3777–3782

Vijayakumar S, Winter K, Sause W, Gallagher MJ, Michalski J, Roach M, Porter A, Bondy M (1998) Prostate-specific antigen levels are higher in African-American than in white patients in a multicenter registration study: results of RTOG 94-12. Int J Radiat Oncol Biol Phys 40:17–25

Zeegers MP, Kiemeney LALM, Nieder AM, Ostrer H (2004) How strong is the association between CAG and GGN repeat length polymorphisms in the androgen receptor gene and prostate cancer risk? Cancer Epidemiol Biomarkers Prev 13:1765–1771