The Polymorphic Exon 1 Androgen Receptor CAG Repeat in Men with a Potential Inherited Predisposition to Prostate Cancer

Ethan M. Lange, Hong Chen, Kristin Brierley, Heather Livermore, Kirk J. Wojno, Carl D. Langefeld, Kenneth Lange, and Kathleen A. Cooney

Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48105 [E. M. L., C. D. L.]; Departments of Internal Medicine [E. M. L., H. C., K. B., H. L., K. A. C.], Surgery [K. A. C.], and Pathology [K. J. W.], University of Michigan Medical School, Ann Arbor, Michigan 48109-0946; and Department of Biomathematics, University of California–Los Angeles School of Medicine, Los Angeles, California 90095 [K. L.]

Abstract

Recent studies have provided epidemiological evidence in support of a possible prostate cancer susceptibility locus on the X chromosome. The androgen receptor (AR) gene, located at Xq11–12, has been implicated as a risk factor for the development of prostate cancer. To examine the potential role of the AR locus in prostate cancer susceptibility, the AR CAG repeat length was measured in 270 Caucasian men with prostate cancer from 133 unrelated families. Each of these families has two or more confirmed cases of prostate cancer occurring in first- and/or second-degree relatives. No evidence for linkage of the AR gene to prostate cancer was observed. We tested for the previously reported association of short CAG alleles with prostate cancer using t tests, Pearson’s χ² tests, and logistic regression; analyses were subsequently repeated to incorporate only men with moderate- to high-grade prostate cancer. No association between AR CAG allele length and prostate cancer was detected when either a subset of unrelated patients or a subset of unrelated patients with moderate- to high-grade cancer was compared with a set of unrelated controls. We failed to detect an association between short AR CAG alleles and early age of prostate cancer diagnosis. Once specific hereditary prostate cancer genes have been identified, future studies can more carefully delineate the potential role of this AR polymorphism as a modifier locus in high-risk families.

Introduction

Several studies have suggested that there may be X-linked prostate cancer susceptibility genes (1–4). The AR gene, located on chromosome Xq11–12, has been considered to be a candidate prostate cancer gene. The gene encodes a transcription factor that binds male sex steroid hormones. Variation of the CAG repeat in exon 1 of the AR gene has also been studied for its possible direct role in prostate cancer causation. Stanford et al. (5) suggested a 3% decrease in prostate cancer risk for each additional AR CAG repeat in a population-based case-control study of middle-aged Caucasian men. Similarly, Giovannucci et al. (6) reported an association between prostate cancer and AR alleles with fewer CAG repeats (relative risk, 1.52) using prostate cancer cases and age-matched controls selected from participants in the Physician’s Health Study. In this latter study, short AR CAG repeat lengths predisposed to higher histological grade and more advanced stage prostate cancer. These associations between short AR CAG alleles and prostate cancer may be a consequence of enhanced transactivation function (7, 8) or increased mRNA levels (9) observed in in vitro experiments using AR genes with fewer CAG repeats.

In 1995, investigators at the University of Michigan initiated the Prostate Cancer Genetics Project with the goal of determining the molecular basis for the inherited predisposition to prostate cancer. We now report the analysis of AR CAG repeat length in 270 Caucasian prostate cancer patients who are participating in this study. We set out to determine whether prostate cancer was linked to the AR gene and whether we could measure an effect of short AR CAG alleles on the occurrence, age of diagnosis, and/or histological grade of prostate cancer in our families.

Materials and Methods

Study Population. The study population consists of 270 Caucasian patients with prostate cancer who are participating in the University of Michigan Prostate Cancer Genetics Project. From this data set, we selected men with prostate cancer from families with two or more confirmed cases of prostate cancer for this present analysis. Written consent was obtained from all participants, and research protocols were approved by the Institutional Review Board at the University of Michigan. The 270 Caucasian prostate cancer patients represent 133 unrelated families. The average age of diagnosis of the 270 men determined by date of prostate biopsy was 64.0 ± 9.5 y (range, 39–90). The diagnosis of prostate cancer was confirmed in 266 of the 270 men by review of pathological records. In the four remaining cases, the diagnosis of prostate cancer was confirmed in a physician’s note, and these cases were, therefore, included in all analyses.

Histological Grade of Prostate Cancer. The pathology records documenting 266 cases of prostate cancer were reviewed without knowledge of the AR CAG repeat length, and 264 cases were categorized into one of three groups: histological grade 1 (G1) with a Gleason sum ≤6 or well-differentiated prostate cancer; histological grade 2 (G2) with a Gleason sum 7 or moderately differentiated prostate cancer; and histological grade 3 (G3) with Gleason sum ≥8 or poorly differentiated prostate cancer. Insufficient information was available in the pathology reports of the remaining two cases to determine
grade. These two cases, along with the four cases without pathological records, were excluded from analyses that included histological grade. In situations where multiple pathology reports were received, priority was given to the grade assigned to the largest volume of resected tumor (e.g., prostatectomy versus biopsy), to the primary diagnosis (instead of recurrence or metastasis), and to the grade assigned by the most experienced genitourinary pathologist in the case of second opinions.

**Determination of AR Exon 1 CAG Trinucleotide Repeat Length.** Genomic DNA was extracted from whole blood using a commercially available kit (Puregene DNA extraction kit; Gentra Systems, Inc., Research Triangle Park, NC). DNA (100 ng) was amplified by two rounds of PCR using nested primers flanking the CAG repeat in exon 1 of the AR gene, with modifications of the protocol of Irvine et al. (10). The CAG repeat lengths were calibrated by comparing PCR product size to the PCR product of a CAG allele, the repeat length of which was determined by direct sequencing.

**CAG Repeat Lengths from a General Caucasian Population.** The allele frequencies of the AR CAG repeat sequence in the general Caucasian population were derived from the studies of Irvine et al. (10) and Stanford et al. (5). The first study reported the AR CAG repeat length from 39 apparently healthy men over the age of 35 yr from Los Angeles, California (10). The second study described similar data from 266 men residing in King County, Washington, who were between the ages of 40–64 yr and had no history of prostate cancer (5). Combining these studies provided AR CAG repeat data from 305 apparently healthy Caucasian men residing in the United States who were over 35 yr of age.

**Linkage Analyses.** To assess the degree of allele sharing among affected relatives at the AR CAG allele size and prostate cancer occurrence. Hence, both one-sided and two-sided statistical tests were computed. Permutation tests (13) using the standard t-statistic were implemented. Because the permutation t test makes no distributional assumptions, derived Ps are more accurate than the Ps of the parametric t test. Empirical Ps cited are based on 1000 or more random permutations of these data.

To analyze the possibility of association between the CAG repeats in the AR gene and prostate cancer, we calculated Pearson’s χ² tests for the relevant 2 × 2 contingency tables (SAS System software; SAS Institute, Inc., Cary, NC). All results are reported as ORs with two-sided 95% CIs. These calculations were conducted using an amended estimator for the OR incorporating a continuity correction as suggested by Gart and Zweifel (14) and Haldane (15). We also performed a logistic regression analysis to determine the OR for a decrease in allele size of one CAG repeat (3 bp). Genotype data were available from more than one affected male in 116 of 133 pedigrees (87%). Due to the potential correlation of affection status and AR CAG repeat length among affected family members, only one family member was included in the hypothesis tests for equal mean allele length and distributional homogeneity. Probands were selected for these analyses in 113 of 116 or 97.4% of families. In the remaining three families, the proband was unaffected; therefore, the first affected family member from which DNA was collected was used. Thus, the preceding hypotheses were tested using 133 unrelated prostate cancer patients and 305 healthy male controls.

Analyses were also performed to explore the relationship between short CAG alleles and high-grade prostate cancer. We used a strategy to select one prostate cancer case with the highest grade from each family that had one or more cases of G2 or G3 prostate cancer (defined as Gleason grade 7–10 or moderate to poorly differentiated cancer, n = 93 families). G3 cases were always selected over G2 cases. If a family had two or more cases of G2 or G3 prostate cancer, one case was randomly chosen for this analysis. Permutation t tests and association tests, as described previously, were implemented on this defined subset of study participants.

To investigate the potential relationship between age of prostate cancer diagnosis and length of the CAG repeat at the AR locus, we implemented the generalized estimation equations (GEE1) approach of Zeger and Liang (16); we assumed Gaussian observations and used the sandwich estimator of the variance to account for the correlation in age of diagnosis among related men. This approach allowed genotype data from all 270 affected men to be incorporated into the analysis. Hypothesis tests were also performed conditioning on histological grade.
Results

NPL analysis using GENEHUNTER version 1.3 revealed a NPL Z-score of −0.73, with a corresponding one-sided P of 0.76. The 24 families with three or more affected men and no evidence of male-to-male disease transmission (4) also failed to show evidence of prostate cancer linkage to the AR gene (NPL Z-score, −0.64; corresponding one-sided P = 0.74).

The allele frequencies for the AR CAG repeat in the probands from our families with two or more cases of cancer (see “Materials and Methods”) are compared graphically to allele frequencies from a control sample in Fig. 1. Permutation t tests were used to compare our population of prostate cancer patients to the control sample. We found no evidence that the mean allele length of the AR CAG fragments in the patient population is smaller than the mean allele length of the AR CAG fragments in the control population (one-sided P = 0.86 and two-sided P = 0.28; Table 1).

To further study the potential relationship between short AR CAG repeats and prostate cancer, we dichotomized the AR CAG repeat, consistent with previous studies that reported a positive association. We found no evidence for an association between prostate cancer and the presence of short alleles at the AR CAG locus using a cutoff of ≤17 versus >17 repeats (17), ≤21 versus >21 repeats (5, 10), or ≤18 versus ≥26 repeats (Ref. 6; Table 2). Results of the logistic regression analysis also revealed that CAG repeat length is not correlated with an increased risk for the development of prostate cancer (OR = 0.96; 95% CI = 0.90–1.03).

We identified 93 unrelated men with moderate-to-poorly differentiated cancer (66 G2 and 27 G3 cases; see “Materials and Methods”). In this group of men, the mean AR CAG allele length was also not significantly different from that of the control population (one-sided P = 0.79 and two-sided P = 0.45; Table 1). Similarly, no difference in CAG allele repeat lengths was detected when this group of men with moderate-to-high grade cancer was compared with controls using three different cutoff values (Table 2). Finally, CAG repeat length was not associated with risk for moderate-to-high-grade prostate cancer using logistic regression (OR = 0.97; 95% CI = 0.90–1.05).

Allele length at the AR CAG polymorphism was found to have no significant effect on the age of diagnosis of prostate cancer (two-sided P = 0.90). The interaction between prostate cancer grade and allele size on age of diagnosis was not found to be significant (two-sided P = 0.57) using a model wherein G1 cases were compared with G2 and G3 cases. Furthermore, no difference in age of diagnosis was detected between the three different grades of prostate cancer after correcting for familial correlation (two-sided P = 0.97).

Discussion

Short repeat lengths of the CAG polymorphism in exon 1 of the AR gene have been hypothesized to predispose to prostate cancer (18). It is speculated that relatively small enhancements of AR activity mediated through increased transactivation and/or mRNA levels promote prostate carcinogenesis over time. However, we failed to detect an association between short AR CAG alleles and prostate cancer incidence when we compared our familial prostate cancer patients to a Caucasian control population. Furthermore, we observed no effect of short CAG alleles and the age of prostate cancer diagnosis nor on the development of high-grade cancer.

There are a number of possible factors that may have contributed to the lack of an observed association between short AR CAG alleles and prostate cancer in our study. The effect of short AR CAG alleles on prostate cancer risk, as determined in two case-control studies, is relatively small, if present at all (relative risk, ≤1.5; Refs. 5 and 6). Lack of a uniform model of analysis also makes these studies difficult to compare. Giovannucci et al. (6) examined CAG repeat length as a continuous variable and also compared men with ≤18 repeats with men with ≥26 repeats. Stanford et al. (5) and Irvine et al. (10) used the median number of 22 to divide their population for analysis. It is unclear whether these studies examined multiple cutoffs and were appropriately corrected for multiple testing. This point is particularly important given the rather weak evidence for association of short AR CAG alleles and prostate cancer incidence, as well as the variety of different models possible for viewing these data.

Population heterogeneity is a potential problem that may be encountered in case-control genetic epidemiology studies. If there are undetected racial/ethnic differences between the cases and unrelated controls, an apparent association between a particular allele and a disease may be confounded. This is a critical concern in studies of prostate cancer, where disease incidence varies dramatically with racial and ethnic background. Our study, as well as the previously reported AR CAG case-control studies, were all subject to the possible effects of genetic heterogeneity.

The patients described here are all participants in the University of Michigan Prostate Cancer Genetics Project; they were selected for this study because of early-onset and/or a
positive family history of prostate cancer. Indeed, 39 of our families (29% of the total of 133 families) fulfilled at least one or more of the proposed clinical criteria for HPC [these criteria are: (a) three or more affected individuals within one nuclear family; (b) affected individuals occurring in three successive generations (maternal or paternal lineage); or (c) a cluster of two or more relatives each affected before the age of 55 yr (19)]. The prostate cancer in these families may be attributable to one or more highly penetrant HPC genes that may mask the relatively modest potential effect of the AR CAG polymorphism. However, Rebbeck et al. (20) recently reported that AR alleles containing very long (≥29) CAG repeats may lead to an earlier age of breast cancer onset in women who also carry a BRCA1 germ line mutation. The role of the AR locus as a modifier of prostate cancer risk in these prostate families can be examined more thoroughly in the future as HPC genes are identified and characterized.

Previous studies have suggested that short AR CAG alleles may predispose to more aggressive forms of prostate cancer, as indicated by high Gleason score tumors and/or advanced stage at diagnosis (6, 17, 21). In our analyses, which incorporated stage, we chose to group all cases up to and including Gleason sum 6, rather than Gleason sum 4, as “well-differentiated” or “favorable.” As we chose to group cases up to Gleason sum 6, rather than Gleason sum 4, as “well-differentiated” or “favorable,” there was also no detectable effect of short CAG alleles on the age of prostate cancer diagnosis or on the development of high-grade cancer in this data set. This is the first comprehensive study of the AR CAG polymorphism in men with early-onset and/or a family history of prostate cancer. Because we could not detect an effect of this polymorphism in our patients, we suggest that the AR gene may play a minor role in the heritable form of this disease. However, as HPC genes are identified, future studies may further delineate the potential role of the AR gene as a modifier locus in prostate cancer families.

Acknowledgments
We thank all of the participants in the University of Michigan Prostate Cancer Genetics Project. In addition, we gratefully acknowledge J. McCarthy, Y. Liu, J. Purakal, and C. Bettis for technical assistance and J. Bailey-Wilson, D. Schot tenfeld, J. Taylor, and K. Piantado for critical review of the manuscript.

References
The Polymorphic Exon 1 Androgen Receptor CAG Repeat in Men with a Potential Inherited Predisposition to Prostate Cancer

Ethan M. Lange, Hong Chen, Kristin Brierley, et al.