Association of Prostate Cancer Risk With Genetic Polymorphisms in Vitamin D Receptor and Androgen Receptor

Sue Ann Ingles, Ronald K. Ross, Mimi C. Yu, Ryan A. Irvine, Giuseppe La Pera, Robert W. Haile, Gerhard A. Coetzee*

Background: Prostate cancer is an increasingly common disease for which there are few well-established risk factors. Family history data suggest a genetic component; however, the majority of prostate cancer cases cannot be explained by a single-gene model. Prostate cell division is influenced by two steroid hormones, testosterone and vitamin D, the action of each being mediated by its respective receptor. The genes for the two receptors are candidate genes in a multigenic model for prostate cancer susceptibility. Purpose: We examined genetic polymorphisms in two steroid receptors, the androgen receptor (AR) and the vitamin D receptor (VDR), in a case-control pilot study of prostate cancer. Methods: Fifty-seven non-Hispanic white case patients with prostate cancer and 169 non-Hispanic white control subjects were genotyped for a previously described microsatellite (CAG repeats) in the AR gene and for a newly discovered poly-A microsatellite in the 3′-untranslated region (3′UTR) of the VDR gene. To compare genotypes with respect to prostate cancer risk, we estimated odds ratios (ORs) by using logistic regression. ORs were also estimated separately for advanced and localized cases of disease. All P values resulted from two-sided tests. Results: Both the AR and the VDR polymorphisms were associated, individually and after mutual adjustment, with prostate cancer. Adjusted ORs (95% confidence intervals [CIs]) for prostate cancer were 2.10 (95% CI = 1.11-3.99) for individuals carrying an AR CAG allele with fewer than 20 repeats versus an allele with 20 or more repeats and 4.61 (95% CI = 1.34-15.82) for individuals carrying at least one long (A18 to A22) VDR poly-A allele versus two short (A14 to A17) poly-A alleles. For both the AR and VDR genes, the at-risk genotypes were more strongly associated with advanced disease than with localized disease. Conclusions: In this pilot study, genetic variation in both the VDR and the AR genes was associated with prostate cancer, and both genes appear to preferentially confer risk for advanced disease. These two genetic risk factors, if confirmed, are among the strongest risk factors yet identified for prostate cancer. Implications: These results are consistent with a multigenic model of prostate cancer susceptibility. On the basis of the joint effect of several genetic loci, one might ultimately be able to construct a risk profile to predict advanced disease, so that men whose disease is unlikely to progress to an advanced stage can possibly be spared aggressive treatment. [J Natl Cancer Inst 1997;89:166-70]

Prostate cancer is an increasingly important medical problem. In 1994, it surpassed breast cancer to become the most frequently diagnosed cancer in the United States (1,2), and the American Cancer Society estimated that 240 000 men will be diagnosed with prostate cancer in 1996 (3). However, little is known about the cause of this disease. Other than age, the most well-established risk factors for prostate cancer are ethnicity and country of residence. Historically, a 50-fold to 100-fold difference in incidence has been reported between the highest and lowest risk populations worldwide (4). African-American men have by far the highest prostate cancer rates in the world. Japanese and Chinese men have the lowest rates, whereas U.S. Caucasians have intermediate rates (5). With the introduction of new detection methods, such as screening for prostate-specific antigen (PSA), it is becoming increasingly difficult to determine the true difference in prostate cancer incidence across populations. Nonetheless, prostate cancer mortality rates, which are less sensitive than prostate cancer incidence rates, demonstrate a threefold to fourfold difference across these same populations. Differences of this magnitude are unlikely to be explained by environmental factors alone.

Family history data have provided suggestive evidence for a genetic component in prostate cancer etiology. First-degree relatives of men with prostate cancer have a twofold to threefold excess risk for prostate cancer; if two first-degree relatives have prostate cancer, the risk is elevated fivefold (6). Although evidence from segregation analysis of multi-case families supports an autosomal dominant mode of transmission, this inherited form is estimated to account for only 9% of all prostate cancers occurring in men by age 85 (7). For the majority of prostate cancer cases, we have proposed a multigenic etiology (8).

We have chosen to examine the genes for two steroid hormone receptors as candidate genes in a multigenic model for prostate cancer susceptibility. Steroid hormone receptors are situated at pivotal points in the pathways that transduce signals from the cellular environment to the nucleus, resulting in transcriptional mod-

*Affiliations of authors: Sue Ann Ingles, R. K. Ross, M. C. Yu, R. W. Haile (Department of Preventive Medicine), R. A. Irvine, G. La Pera, G. A. Coetzee (Department of Urology), Norris Comprehensive Cancer Center, University of Southern California School of Medicine, Los Angeles.

Correspondence to: Sue Ann Ingles, Dr.P.H., University of Southern California/Norris Comprehensive Cancer Center, 1441 Eastlake Ave., MS#44, Rm. 4415, P.O. Box 33800, Los Angeles, CA 90033-0800.

See “Notes” following “References.”
ulation of genes involved in cell proliferation and differentiation. Since steroid receptors amplify these signals, a small change in receptor number or function could potentially have a large impact on cell proliferation and/or differentiation.

Testosterone is the principal steroid hormone in prostate epithelial cells. It is primarily responsible for driving cell division in the prostate. For androgen activity to manifest in the prostate, however, 5α-dihydrotestosterone (DHT), the active form of testosterone, must bind to the androgen receptor (AR) and translocate to the nucleus, where transactivation of target genes is mediated by the AR–DHT complex. The transactivation region of the AR gene is known to be polymorphic, and one polymorphism, a CAG repeat in exon 1, is the cause of X-linked spinal and bulbar muscular atrophy (Kennedy’s disease) (9). Patients with this disease have ARs with more than double the average number of CAG repeats. These mutant receptors have a reduced transactivation activity even in the presence of normal androgen binding (10). Moreover, within the normal range of CAG repeats (i.e., nine to 30 CAG repeats), the AR transactivation level in vitro is inversely related to the number of CAG repeats (11). Given this evidence, we have hypothesized and have subsequently provided case-control data (12) that the CAG repeat might be associated with the development of prostate cancer (13). In this study, we expand on these results by comparing the non-Hispanic white case patients from the previous study (12) with a much larger control group while we jointly estimated the effect of a second hormone receptor locus.

A second steroid hormone known to influence cell division in prostate cells is 1,25(OH)2D3, the active form of vitamin D. 1,25(OH)2D3 has been shown to cause both antiproliferative and prodifferentiation responses in prostate cancer cell lines (14,15). Although the specific genes responsible for these vitamin D-mediated effects have not been identified, a number of cancer-associated genes have been shown to be modulated directly or indirectly by vitamin D (16,17). Among these genes, the cdk inhibitor p21 has been shown to be directly induced by binding of the vitamin D3–vitamin D receptor (VDR) complex to vitamin D response elements in the p21 promoter (18).

The human VDR is known to have polymorphisms that result in individual variation in the circulating level of a VDR-regulated protein, osteocalcin (19). Although the exact sequence elements responsible for functional variation are not known, the alleles can be assayed by use of restriction polymorphisms (EcoRV, Bsm I, Apa I, and Taq I) that lie in the region from exon 8 to the 3′-untranslated region (3′UTR). The 3′UTRs associated with the two most common haplotypes exhibit a twofold difference in VDR gene expression in a reporter gene assay, and sequencing of these two common haplotypes revealed 13 differences (including point mutations, insertions, deletions, a 3- or 4-base-pair substitution, and a poly-A length polymorphism) (20). Presumably, one or more of these sequence variations are responsible for differences in translation efficiency or in messenger RNA (mRNA) stability, resulting in individual variation in VDR expression. We have developed an assay for one of these 13 candidate sequences, the poly-A polymorphism. In this study, we assess the possible association of this VDR polymorphism and the AR CAG polymorphism (described earlier) with prostate cancer risk.

Subjects and Methods

Subjects

All case patients were non-Hispanic white men previously described in a study of the AR (12). Briefly, prostate cancer cases diagnosed in 1991 through 1992 were identified by the Cancer Surveillance, Epidemiology, and End Results (SEER) cancer registry of Los Angeles County. In early 1994, five hundred fifty-nine (53%) of 1062 eligible case patients responded to a mailed questionnaire that requested detailed information on family history of cancer. The first 68 incident case patients who gave written informed consent to participate in a family-based study of prostate cancer were included in the pilot study described here. Fifty-seven of these case patients were non-Hispanic whites. The three other ethnic groups represented were African-American, Egyptian, and Chinese-American, but they did not contribute enough cases for meaningful statistical analysis; thus, 11 individuals were excluded, leaving 57 case patients in the study. Nine of these 57 case patients (16%) had one or more first-degree relatives (father, brother, or son) with prostate cancer. The age at diagnosis ranged from 51 years to 68 years, with a mean age of 57.8 years and a median age of 58 years.

One hundred seventy-one control subjects (control subject–case patient ratio = 3:1) were selected from non-Hispanic white males who were in an ongoing population-based case–control study of bladder cancer in Los Angeles County. These control subjects were selected systemically from the neighborhoods in which the bladder cancer patients resided. Using the house of each case patient as a reference point and proceeding in a systematic and invariable sequence, we canvassed up to 150 residential units to identify two control subjects who could be matched to the case patients by sex, year of birth (within 5 years), and race (Hispanic white, non-Hispanic white, or black). The control subjects reported on here were selected to be comparable to the prostate cancer patients in age and in socioeconomic status. The mean age of the control subjects was 58.2 years. Of these 171 control subjects, two were excluded because of polymerase chain reaction (PCR) amplification failure of one or both of the markers, leaving 169 control subjects in the study. Written informed consent was obtained from all subjects, and the investigations were approved by the appropriate institutional review board.

Genotyping of Polymorphisms

Two polymorphic microsatellites were genotyped: a CAG microsatellite in the AR gene and a poly-A microsatellite in the VDR gene. Both microsatellites are length polymorphisms, with alleles varying in the number of repeated units (CAG repeats or adenosine residues [As]). Individual alleles will be identified by the number of repeats that they contain, which we will henceforth call “allele size.” The DNA was extracted from white blood cell suspensions obtained as buffy coats from heparinized blood by ficoll-hypaque separa-

Report 167

Journal of the National Cancer Institute, Vol. 89, No. 2, January 15, 1997

167
lowest tertile includes 55% of African-Americans but only 21% of Asians, whereas the highest tertile includes 51% of Asians but only 25% of African-Americans (12).

Unconditional logistic regression was used to estimate unadjusted (univariate) odds ratios (ORs) and to estimate ORs for each of the two loci while adjusting for the other locus. A test of interaction was performed by adding an interaction term to the logistic model and computing the likelihood ratio statistic (22). Where indicated, ORs and confidence intervals (CIs) for subgroups of case patients were estimated by use of exact distributions with a mid-P correction (23) as implemented by the program StatXact Turbo 2.04 (Cytel Software Corp., Cambridge, MA).

Because the results of significance tests may depend on the choice of cut points for a continuous variable (24) and because cut points were arbitrarily chosen in the case of the CAG polymorphism, differences among the CAG allele size distributions were also tested for significance by use of the Wilcoxon test (25). All P values resulted from two-sided tests.

Results

AR Polymorphisms

For the non-Hispanic white subjects in this study, those in the lowest tertile (<20 CAG repeats) had an approximately two-fold increase in risk compared with those in either the highest tertile (≥22 CAG repeats) or those in the middle tertile (20-21 CAG repeats) (Table 1). Since risk for the middle and highest tertiles did not differ substantially, allele size was dichotomized as long (≥20 CAG repeats) and short (<20 CAG repeats) in the logistic regression analyses.

VDR Polymorphisms

For the VDR poly-A microsatellite, all allele sizes from A14 to A22 were observed. The distribution of allele sizes was bimodal, with alleles A15 to A17 being rare (Fig. 1). A single poly-A long allele (A18 to A22) was associated with a fourfold to fivefold increase in risk. Homozygosity for long poly-A alleles did not appear to increase risk substantially (Table 1); therefore, the genotypes were grouped as no long alleles (i.e., both short alleles) and one or more long alleles in the remaining analyses.

Stratification by Stage of Disease

Thirty-one case patients (54%) had disease localized to the prostate, and 26 (46%) had advanced disease (defined as a tumor invading and extending beyond the prostatic capsule and/or extending into adjacent tissue or involving regional lymph nodes or distant metastatic sites, SEER 1995 clinical and pathologic extent of disease codes 41-85). Both the AR short alleles (<20 CAG repeats) and the VDR poly-A genotypes with at least one or both long alleles were more strongly associated with advanced disease than with localized disease (Table 2). Moreover, the genotype–disease associations were statistically significant for advanced disease but not for localized disease. Because cut points for the CAG polymorphism were arbitrarily chosen, we verified these results by comparing CAG allele size distributions by use of a nonparametric test that did not depend on cut points. Compared with values for control subjects, the AR allele size distribution (Fig. 2) was found to be shifted toward smaller values for patients with advanced disease (P = .05) but not for patients with localized disease (P = .73).

Multivariate Analysis

Results of the analysis of the association of prostate cancer risk with each of the two steroid hormone receptors mutually adjusted for the genotype at the other locus are presented in Table 2. Adjusting one locus for the other did not substantially alter the effect estimates. An AR CAG allele of length less than 20 was associated with an approximate doubling of risk, whereas the presence of one or more long VDR alleles was associated with a fourfold to fivefold increase in risk.

We also tested whether the two loci might interact in a multiplicative manner, such that individuals with high-risk alleles at both loci might be at a higher (or lower) risk than the risk that would be predicted by the combined effects of the individual loci. We found that the interaction was not statistically significant (P = .51), indicating no interaction between the two loci. Given the number of subjects in this study, however, the power to detect a moderate-sized interaction was low (e.g., 30% power to detect ORinteraction = 2.5).

Discussion

We have demonstrated that polymorphisms in two steroid hormone receptor genes are associated with prostate cancer in non-Hispanic white males. The simplest hypothesis that might explain these findings is that the steady-state expression of these receptor alleles or the functional activity of their products differs to the extent that they impose differences in disease risk.

The most attractive hypothesis to explain the effect of AR short alleles is that they impose a quantitatively higher transactivation activity on the receptor, as was shown to be the case in in vitro experiments (13). Higher transactivation activity may result in increased proliferative activity in prostate cancer cells; however, the downstream genes involved have not been identified. Androgens are required for prostate cancer growth; thus, higher transactivation activity might increase the risk of prostate cancer, particularly the risk of advanced disease.

In contrast to the AR CAG repeats, the poly-A microsatellite in the 3′UTR of the VDR gene may simply mark the relevant locus, although we have not excluded a possible direct effect on mRNA function, perhaps by differentially binding transacting factors like poly(A)-binding protein (26). Alternatively, the VDR polymorphism may be in linkage disequilibrium with 3′UTR sequence elements that affect
mRNA function (e.g., AU-rich elements that, in response to metabolic signals, bind specific proteins that destabilize the mRNA). Whatever the mechanism, it is possible that the VDR 3′UTRs marked by the long alleles cause decreased VDR expression, leading to decreased transcriptional control of VDR-regulated genes. Dysregulation of metastasis-related genes such as the VDR-regulated fibronectin gene (27) might increase risk of advanced disease.

In this report, the genetic analyses were conducted on samples collected for purposes other than those of interest in the present study. Therefore, the comparability of the sources for the prostate cancer patients versus the healthy controls must be scrutinized carefully. The prostate cancer patients were drawn from a population-based source and likely represent a reasonably random sampling of non-Hispanic white patients under 70, whose prostate cancer was diagnosed in Los Angeles during 1991 and 1992; however, because this is a young group, the results may not be generalizable to all prostate cancer patients. The control subjects were derived from a population-based source also and were selected via a carefully followed algorithm of door-to-door census taking by trained research assistants and were then further selected to match the age, race, and socioeconomic status of the prostate cancer patients. Although we believe that this procedure resulted in good comparability of source populations, the case patients and control subjects were nonetheless ascertained for separate studies, so it is especially important that our results be confirmed by use of more standard epidemiologic approaches.

If confirmed, the two genetic risk factors that we have identified (VDR and AR alleles) are among the strongest risk factors yet identified for prostate cancer. However, the patients in this study did not have a striking family history of prostate cancer (only 16% having a first-degree relative with the disease). This observation is consistent with a multigenic model of prostate cancer etiology. In such a model, no single gene is sufficient to produce a mendelian pattern of disease segregation; rather, disease risk is influenced by several genes and possibly by gene–gene and gene–environment interactions. Since we found no evidence for gene–gene interaction (albeit with limited statistical power) and since the effects of allelic variation on risk remained essentially unchanged when we adjusted each locus for variation at the other locus, the AR and VDR polymorphisms appear to contribute independently to prostate cancer risk.

Both AR and VDR polymorphisms appear to preferentially confer risk for advanced disease rather than for localized disease; however, larger studies are clearly needed to confirm these results. Predic-

### Table 2

<table>
<thead>
<tr>
<th>AR CAG repeats*</th>
<th>Odds ratio (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
</tr>
<tr>
<td>Long (%)</td>
<td>Short (%)</td>
</tr>
<tr>
<td>Control subjects</td>
<td></td>
</tr>
<tr>
<td>All case patients</td>
<td>124 (73)</td>
</tr>
<tr>
<td>Patients with localized disease</td>
<td>33 (58)</td>
</tr>
<tr>
<td>Patients with advanced disease</td>
<td>19 (61)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VDR poly-A‡</th>
<th>Odds ratio (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
</tr>
<tr>
<td>SS (%)</td>
<td>SL/LL (%)</td>
</tr>
<tr>
<td>Control subjects</td>
<td>33 (20)</td>
</tr>
<tr>
<td>All case patients</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Patients with localized disease</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Patients with advanced disease</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

*AR CAG allele size: short = <20 CAG repeats; long = ≥20 CAG repeats.
†Odds ratios mutually adjusted for variation at the other locus.
‡VDR poly-A allele size: S = A_{14} to A_{17}, and L = A_{18} to A_{24}.
§Confidence limits estimated by exact method.
\[\text{Odds ratio cannot be estimated because of empty cell; lower confidence limit estimated by exact method.}\]
tors of disease progression are needed to address the dilemma of whether or not to aggressively treat all men who have elevated PSA levels when only a fraction of these men will develop advanced disease. On the basis of the joint effect of several loci, one might ultimately be able to construct a risk profile that could predict advanced disease and could allow for more meaningful decision-making regarding optimal treatment strategies.

References


Notes

1Editor's note: SEER is a set of geographically defined, population-based central tumor registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Each registry annually submits its cases to the NCI on a computer tape. These computer tapes are then edited by the NCI and made available for analysis.

Supported in part by Public Health Service grants CA17054, CA65726, and UCLA CNRU CA42710 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services; and by the Elsa U. Pardee Foundation.

We gratefully acknowledge Lucy Xia and Wu Zhang for their expert technical assistance.

Manuscript received May 29, 1996; revised September 9, 1996; accepted October 10, 1996.