Genetic Variation at 8q24 as a Susceptibility Factor for Prostate Cancer: Definitive Results from Epidemiologic Studies?

□□*Commentary on Wang et al., p. 2944 and Schumacher et al., p. 2951*

Elizabeth A. Platz

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health and the James Buchanan Brady Urological Institute and the Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins Medical Institutions, Baltimore, Maryland

In this issue of *Cancer Research*, two epidemiologic studies investigating the association between genetic variants on chromosome 8q24 and prostate cancer are published. The purpose of this editorial is to comment on the importance and implications of these findings and to explain the rationale of the journal for publishing not just one but two epidemiologic studies on 8q24 and prostate cancer.

The long arm of chromosome 8, including 8q24, has been the focus of research in prostate cancer for some time. It consistently shows gains in prostate adenocarcinomas, especially in metastatic cases (1), and some groups report that 8q gains are predictive of poor outcome (2, 3). Last year, the *8q24* locus was first observed to be linked to prostate cancer in a large Icelandic family study in a genome-wide linkage scan conducted by Amundadottir et al. (4) at deCODE genetics. When the same group evaluated a large number of microsatellite and insertion/deletion markers around 8q24, they found that the -8 allele at marker DG8S737, an AC dinucleotide repeat located at 8q24.21, was the one that was most strongly associated with prostate cancer in a set of Icelandic cases and population-based controls. Additional work narrowed the area to a single nucleotide polymorphism (SNP; rs1447295), which was found to be associated with a 50% increased risk of prostate cancer, along with a 60% increased risk for the -8 allele, in pooled case-control studies of White men from Sweden and the United States. The work conducted by deCODE genetics was impressive for two reasons: (a) the enormous scope of the undertaking, which encompassed genome-wide linkage scanning, SNP discovery, genotyping of a large number of cases and controls for a large number of markers, and also a search for RNA sequences transcribed from this region in prostate and other tissues; and (b) the logical progression of scientific discovery and testing, extending from a family-based linkage study to a single population-based case-control study to a replication case-control study in the same country to "validation" case-control studies in populations of other white ethnicities and finally, to a case-control study in another racial group: African-Americans who are disproportionately affected by the burden of prostate cancer. When investigated in an African-American population, a 60% higher odds for the -8 allele was observed but only a 15% higher for the A allele (4). For both variants, the ORs were only marginally stronger for high-grade prostate cancer (Gleason sum 7+ on biopsy) than for low-grade disease. The results of the work conducted by deCODE genetics are notable for the consistency of the findings from study to study and among racial and ethnic groups.

A second collaborative group independently pointed to 8q24 using a whole-genome admixture scan of ~1,600 African-American men with prostate cancer from seven studies, including the Multiethnic Cohort (5). Ancestral admixture scanning involves comparing regions of the genome that contain more or less than the genome-wide average admixture of ancestral DNA markers within cases. The investigators selected genetic markers for which the prevalences differed most greatly between West Africans and Whites of European ancestry. Carrying 8q24 markers that were common for West African ancestry was associated with a higher risk of prostate cancer, especially at a younger age at onset. The investigators also evaluated the -8 allele at DG8S737 and the A allele at rs1447295 in the African-American prostate cancer cases who were diagnosed when they were younger than 72 years old and age-matched controls sampled from the seven studies. They noted no association for either variant after taking into account possible differences between the cases and controls in the extent of West African ancestry. However, when Freedman et al. (5) conducted a case-control study nested in the prospective Multiethnic Cohort, they noted ~40% to 50% higher odds of prostate cancer for the A allele at rs1447295 among Japanese Americans, Latino Americans, and European Americans; the association was stronger in Native Hawaiians. The association for the SNP did not differ statistically significantly by grade (Gleason 8+ versus <8).

As in the deCODE study, this work done by Freedman et al. (5) was impressive in the breadth of the research approach, using both admixture mapping and a prospective epidemiologic study, and in the extent of the effort involved. What was particularly notable about their results was again the consistency of findings: (*a*) in the admixture study, the logarithm of odds stayed high after omitting the cases from each study one by one and (*b*) in the association studies, the odds ratio (OR) for the SNP was elevated in each non–African-American racial/ethnic group.

The consistency of the work published by Amundadottir et al. (4) and Freedman et al. (5) provides some evidence that genetic variation at 8q24 is a risk factor for prostate cancer. However, epidemiologic studies need to be interpreted with caution, as there is always room for selection bias, observation bias, chance, and confounding to explain a finding. We believe that the additional evidence from the National Cancer Institute's Cohort Consortium's series of nested case-control studies by Schumacher et al. (6) and the Mayo Clinic's case-control study by Wang et al. (7) strongly confirms the importance of the *8q24* locus as a susceptibility allele for prostate cancer. Furthermore, these two new studies bring additional information about the nature of the association.

Editor's Note: Dr. Platz has published articles with some of the authors of the Schumacher et al. study but has not been involved in any of the work on 8q24 reported by them.

Requests for reprints: Elizabeth A. Platz, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Room E6143, 615 North Wolfe Street, Baltimore, MD 21205. Phone: 410-614-9674; Fax: 410-614-2632; E-mail: eplatz@jhsph.edu.

doi:10.1158/0008-5472.CAN-07-0737

Schumacher et al. (6) report findings for the variant A allele at rs1447295 from the Cohort Consortium; the microsatellite repeat was not studied. The Consortium consists of seven cohorts, among which more than 6,600 prostate cases have been diagnosed. Using a nested case-control design within each cohort, one of the two gold standard approaches for the efficient study of genetic and molecular markers in relation to disease risk, the investigators in total sampled more than 7,300 men without prostate cancer as age-matched controls from the at risk person-time. In this very well-designed and conducted epidemiologic study, the investigators used a stringent criterion for statistical significance and report confidence intervals that are 99% rather than 95%, which equates to a type I error of 1% (significant if $P < 0.01$) rather than 5%.

Among Whites, Schumacher et al. (6) observed that, compared with men with two copies of the C (major) allele, men with two AA alleles at rs1447295 had an ~90% higher risk of prostate cancer, and those with one A allele had a 30% higher risk. Importantly, Schumacher et al. (6) also evaluated this locus in relation to risk of breast cancer among 2,600 women with breast cancer and 3,100 controls from the four cohorts in the Cohort Consortium and found that, in contrast to prostate cancer, there was no association at all between the A allele at rs1447295 and breast cancer risk in Caucasian women, whether *in situ*, localized, or metastatic disease.

What is most remarkable in this work is the consistency of the association across each of the seven cohorts: comparing the AA and CC genotypes, the minimum OR was 1.39 among men in the prostate-specific antigen (PSA) screening arm of a large cancer screening trial (Prostate, Lung, Colorectal, and Ovarian Trial) to a maximum OR of 2.77 in one of the American Cancer Society's large cohorts. Note that the associations were not statistically significant for every cohort, but the direction of the association was the same and the heterozygotes had an odds ratio that was intermediate between the homozygous wild-type and the homozygous variant in each of the cohorts. The association was present irrespective of the nature of the case, early versus late or fatal and low versus high grade, although arguably (but not statistically), the association was slightly stronger for cases with a worse prognosis or who died from prostate cancer. There was no difference by family history, defined by the self-report of a father or brother with prostate cancer.

Given the prospective approach and enormous sample size of the Cohort Consortium study of 8q24 and prostate cancer (6), why did *Cancer Research* also publish the study by Wang et al. (7), which used the retrospective approach and had a considerably smaller sample size? The strength of the study by Wang et al. (7) is that it included an assessment of the association for familial prostate cancer cases, which were ascertained from known pedigrees with multiple affecteds. Even in a very large case-control study nested within a cohort, the number of men with familial prostate cancer would be small.

Wang et al. (7) conducted a study of 1,100 Caucasian men with prostate cancer and 595 population-based controls to investigate the association of the DG8S737 microsatellite repeat and the rs1447295 SNP with prostate cancer. The investigators enriched the prostate cancer cases with familial disease, defined as cases from families with at least three affected members, and aggressive disease, defined as Gleason 8+ irrespective of family history. The sporadic cases were men who did not have a family history of prostate cancer and controls were men recruited from the surrounding community who did not have prostate cancer on urologic work up.

In this study, the controls were unselected with respect to family history, so their *8q24* allele prevalences should be undistorted relative to the source population. In sampling both types of cases, the investigators induced selection bias (i.e., the case or control sampling method is dependent on exposure status), which they used to advantage. Given that genetic variation at 8q24 is inherited and if that genetic variation is associated with prostate cancer, then a family history of prostate cancer should be correlated with 8q24 genotype, at least weakly. By restricting sporadic cases to men who did not have a family history of prostate cancer, the investigators would, in theory, have reduced the prevalence of the "causal" allele in those cases (i.e., the included cases would be those with a nongenetic etiology). By restricting familial cases to men who had 3+ affected family members, the investigators would, in theory, have increased the prevalence of the "causal" allele in those cases. Given Wang et al.'s (7) sampling strategy, we would expect to observe no association or an underestimated association between 8q24 and prostate cancer when the sporadic cases were compared with the population-based controls and an enhanced association when the familial cases were compared with the population-based controls. These biased sampling strategies can have usefulness in genetic epidemiology studies. Because the familial case group should be enriched for inherited causal alleles compared with controls who are unselected for family history, a smaller sample size is required to detect modest to moderate associations than when the case group is not enriched for family history. In light of an association for total prostate cancer and/or familial prostate cancer, an attenuated or null result for the sporadic case group may provide complementary evidence for the genetic basis for an association.

Indeed, Wang et al. (7) observed essentially no association between carrying at least one A allele at rs1447295 or -8 allele at DG8S737 and sporadic prostate cancer. We would expect to observe a lower allele prevalence in the sporadic cases than in cases unselected for family history. When examining the two variants in the sporadic cases, the allele prevalences were not inconsistent with those observed in the cases unselected for family history in the other studies, although they were at the low end of the range. Overall, we cannot be sure whether the lack of a clear association between this locus and sporadic disease is because of the family history restriction or another reason. Both the A and the -8 alleles were positively associated with prostate cancer when familial cases were compared with controls, and, indeed, the OR for carrying at least one variant allele for the SNP for the familial cases versus controls was greater (OR = 1.93) than the ORs observed for carrying at least one variant allele in the Cohort Consortium (OR $\approx$ 1.4), in which the cases and controls were unselected with respect to family history. Although the OR is likely overestimated, nevertheless, the association is present in the correct direction and thus gives further evidence for this locus in prostate cancer.

For aggressive disease, which was not sampled with respect to family history, there was a positive association for the A allele but not for the -8 allele. Instead the -10 allele seemed to be positively associated with aggressive disease. It would be interesting to revisit their results in detail in each of the other studies.

In summary, we have never observed such consistency for prostate cancer genetics, not among linkage studies, not among association studies, and certainly not when combining all approaches together. Unless the bias is exactly the same irrespective of study design [linkage, admixture, case-control, and (prospective) nested case-control] and population (country,

race/ethnicity, and prevalence of PSA screening), which is extremely unlikely, then we can conclude with some confidence that genetic variation at 8q24 is important for prostate cancer risk in all populations. The locus seems to be associated with all stages and grades of prostate cancer, albeit slightly more strongly with cases with a worse prognosis. Perhaps genetic variation at this locus influences the development of prostate cancer or affects the detectability of prostate cancer. The prevalence of the A allele at rs1447295 ranged from 8% to 17% in White controls across all of the studies (4–6) and was 31% in African-American controls in both the Flint Men's Health Study and the Multiethnic Cohort (4, 5). Because of differences in allele prevalences, this locus may explain more of the incidence of prostate cancer in some populations, like African-American men, than in others, a point highlighted in three of the studies (4–6).

We still do not know the exact location of the causal allele or the biological mechanism underlying the 8q24-prostate cancer association. Neither the -8 allele of DG8S737 nor the A allele of rs1447295 is likely the causal variant. The two alleles are moderately correlated (4), and in two of the studies considering the two alleles jointly (7), the odds ratio was higher for both alleles than for either one alone. If one had been causal, the extra information provided by the other marker would not have changed the estimate from the main effect to the joint effect. Further, the now narrowed region on 8q24 does not seem to encode any known gene; deCODE genetics looked for transcripts and found a couple expressed in normal prostate and in prostate cancer cell lines but neither was homologous to known genes. *c-MYC* is the closest known cancer-associated gene, but additional work done by deCODE genetics showed that their findings were not explained by variation in *c-MYC* (4). The lack of information on the biological mechanism underlying this consistent association between 8q24 and prostate cancer does not rule out the importance of 8q24 in the etiology of prostate cancer. Certainly, a mechanism is needed to determine if a preventive or intervention strategy is possible. Why this locus is so consistently associated with all "types" of prostate cancer but was clearly not associated with breast cancer overall or by invasiveness in any of the four cohorts that included women in the Cohort Consortium is a mystery that deserves attention. We expect that several basic science reports will be published over the next few months that will move us closer to knowing the causal locus and perhaps a novel mechanism for the association between the *8q24* locus and risk of prostate cancer.

Acknowledgments

Received 2/22/2007; accepted 2/22/2007.

References

1. Nupponen N, Visakorpi T. Molecular biology of progression of prostate cancer. Eur Urol 1999;35:351–4.
2. Ribeiro FR, Jeronimo C, Henrique R, et al. 8q gain is an independent predictor of poor survival in diagnostic needle biopsies from prostate cancer suspects. Clin Cancer Res 2006;12:3961–70.
3. van Duin M, van Marion R, Vissers K, et al. High-resolution array comparative genomic hybridization of chromosome arm 8q: evaluation of genetic progression markers for prostate cancer. Genes Chromosomes Cancer 2005;44:438–49.
4. Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006;38:652–8.
5. Freedman ML, Haiman CA, Patterson N, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proc Natl Acad Sci U S A 2006;103:14068–73.
6. Schumacher R, Feigelson H, Cox D, et al. A common 8q24 variant in prostate and breast cancer from a large nested case-control study. Cancer Res 2007;67:2951–6.
7. Wang L, McDonnell S, Slusser J, et al. Two common chromosome 8q24 variants are associated with increased risk for prostate cancer. Cancer Res 2007;67:2944–50.