AD_____

Award Number: W81XWH-06-1-0181

TITLE: Analysis of Ethnic Admixture in Prostate Cancer

PRINCIPAL INVESTIGATOR: Cathryn H. Bock, Ph.D.

CONTRACTING ORGANIZATION: Wayne State University
Detroit, Michigan 48202-3622

REPORT DATE: December 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-12-2006	Annual	30 Nov 2005 – 29 Nov 2006

4. TITLE AND SUBTITLE

Analysis of Ethnic Admixture in Prostate Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0181

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Cathryn H. Bock, Ph.D.

E-Mail: bockc@med.wayne.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Wayne State University
Detroit, Michigan 48202-3622

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Evidence for a genetic component to prostate cancer is strong, however few genes have been identified, and most of the genetic risk remains undefined. To date, multiple traditional genome scans and linkage analyses have been performed, and several susceptibility loci and candidate genes have been identified. The goal of this research proposal is to use a novel approach to gene discovery, admixture mapping, to identify potential prostate cancer susceptibility genes in a group of African American men. Admixture mapping has greater power to detect genetic effects than traditional genome linkage scans. Recently, Freedman et al. published results from an admixture mapping study of prostate cancer in 1,597 African American men which detected a susceptibility region on chromosome 8q24. In the current study, approximately 900 samples from 2 case-control study of prostate cancer are being genotyped for ancestry informative markers across the genome, using a similar marker panel to that used by Freedman et al. Regions showing strong linkage using the admixture mapping approach will be followed by future studies using fine mapping with a denser set of informative markers in the regions of interest and candidate gene studies. After 12 months, the project is on time with completion of the targeted tasks outlined in the Statement of Work for the project's first year, and is on track to complete all tasks within the next 12 months, as planned.

15. SUBJECT TERMS
Prostate cancer, admixture mapping

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	23	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

INTRODUCTION:

Evidence for a genetic component to prostate cancer is strong, however few genes have been identified, and most of the genetic risk remains undefined. To date, multiple traditional genome scans and linkage analyses have been performed, and several susceptibility loci and candidate genes have been identified, including *HPC1*, *HPCX*, *HPC20*, *CAPB*, *PCAP*, *RNASEL*, *HPC2*/*ELAC2*, and *MSR1*. Traditional genome scans using information from prostate cancer families, however, have generally not included enough African American families to provide adequate statistical power to detect linkage. The goal of this research proposal is to use a novel approach to gene discovery, admixture mapping, to identify potential prostate cancer susceptibility genes in a group of African American men. Admixture mapping has greater power to detect genetic effects than traditional genome linkage scans. Recently, Freedman et al. published results from an admixture mapping study of prostate cancer in 1,597 African American men which detected a susceptibility region on chromosome 8q24 (manuscript included in Appendix) (1). In the current study, approximately 900 samples from 2 case-control study of prostate cancer are being genotyped for ancestry informative markers across the genome, using a similar marker panel to that used by Freedman et al. The admixture mapping analyses will be performed using ADMIXMAP and ANCESTRYMAP statistical programs. Regions showing strong linkage using the admixture mapping approach will be followed by future studies using fine mapping with a denser set of informative markers in the regions of interest and candidate gene studies.

BODY:

After 12 months, the project is on time with completion of the targeted tasks outlined in the Statement of Work for the project's first year, and is on track to complete all tasks within the next 12 months, as planned. Almost all of Task 1 is completed, and Task 2 activities have been started. Task 3 activities are not slated to commence until month 20. The details of progress within each Task are as follows.

Task 1. To obtain genotype information for all study subjects (Months 1-18):

- *a. Prepare batches of DNA and ship to ParAllele, starting with samples from controls (Months 1-17).*
- *b. ParAllele to perform Genotyping and transmit results to Dr. Bock (Months 2-18).*

DNA samples for approximately 530 cases and 380 controls were prepared for genotyping, and shipped to the laboratory. In addition to the approximately 250 cases and 100 controls from HFHS, we also included approximately 280 cases and 280 controls from Dr. Rick Kittles' prostate cancer case control study (described in Bonilla et al., (2) included in the appendix) (with IRB approval), thereby more than doubling our sample size. For quality control, DNA samples from 30 CEPH individuals were included so that their genotype results could be compared with those publicly available through HapMap. Because ParAllele was out of business when we were ready to genotype, we used a panel of 1536 ancestry informative SNPs developed by David Reich at the Broad Institute for use on the Illumina BeadStation platform. Earlier versions of

this panel were used in the Freedman et al. prostate cancer admixture mapping study (1). This panel has very high reliability and success rate in Dr. Reich's lab, and after running 80% of the samples in the Wayne State University Genomics Core Laboratory, there were only two samples that were not typable on their first run, and >95% of the markers had excellent results. The two samples that did not provide reliable results will be re-run with the last batch of samples in January, 2007. It is anticipated that all of the genotyping will be completed and the results provided to the PI in January, 2007. Thus, all of Task 1 should be completed by the end of January, 2007, within the timeframe outlined in the Statement of Work.

Task 2. To identify candidate prostate cancer susceptibility loci using mapping by admixture linkage (MALD) (Months 1-22).

a. Set up database and preliminary ADMIXMAP program (Months 1-6).

b. Perform preliminary analyses and refine ADMIXMAP program (Months 6-18).

c. Calculate final LOD scores and 95% confidence intervals for regions that show possible linkage (Months 18-22).

d. Where necessary, extend the score test and likelihood ratio tests in ADMIXMAP to test for gene-environment interactions. (Months 20-22).

The ADMIXMAP software was successfully set up on a local computer, and the PI successfully ran a test data set through the program. We will begin performing preliminary analyses in late January or early February, 2007, within the timeframe outlined in the Statement of Work. We will also send the genotype data to Dr. David Reich, the senior author of the recent report on prostate cancer admixture mapping results (1). He is an expert in this area, and has agreed to run our data through his ANCESTRYMAP program, which will allow us to directly compare our results with those he published. We anticipate that the analyses outlined in Task 2 will be completed according to the stated times.

Task 3. Final Analyses and Report Writing, Months 20-24:

a. A final report describing the mapping findings and any gene-environment interactions will be prepared (months 20-24).

We anticipate accomplishing Task 3 on schedule, after completing Task 2.

KEY RESEARCH ACCOMPLISHMENTS:

- Prostate cancer case and control DNA samples from African American men from two case control studies were prepared to for genotyping and shipped to the lab for genotyping. The number of samples for genotyping was more than double the expected number of originally anticipated samples.
- An improved panel of 1536 ancestry informative markers was identified and purchased for use in genotyping the samples.

- Genotyping of all samples is 80% complete, and the results have been provided to the PI.
- Genotyping is expected to be 100% complete and all final results transmitted to the PI by Jan 31, 2007.
- ADMIXMAP Software (http://www.ucd.ie/genepi/admixmap/index.html) was installed successfully, and test data was successfully run through the program.
- A new collaborator, David Reich, joined the project and has agreed to provide his expertise in admixture mapping and also to run the data through his admixture mapping program, ANCESTRYMAP (3). IRB approval was obtained to provide him with the genotype data.
- Initial analyses in ADMIXMAP and ANCESTRYMAP will be performed beginning in late January or early February, 2007 to identify potential candidate regions for prostate cancer susceptibility genes.

REPORTABLE OUTCOMES:

A database of genotype information on 80% of the subjects has been established, however the project has not yet reached the point in its timeline when any reportable outcomes regarding prostate cancer risk can be stated. These will be forthcoming as the data is analyzed early in 2007.

CONCLUSION:

This study is on track with its Statement of Work goals and timeline; the sample size is more than double the anticipated number, genotyping is almost completed with high quality of results obtained to date, and the analysis programs are in place to be run. Because genotyping will need to be completed before definitive analyses can be run, there are not yet specific conclusions regarding prostate cancer susceptibility loci available from this project.

REFERENCES:

1. Freedman ML, Haiman CA, Patterson N, McDonald GJ, Tandon A, Waliszewska A, Penney K, Steen RG, Ardlie K, John EM, Oakley-Girvan I, Whittemore AS, Cooney KA, Ingles SA, Altshuler D, Henderson BE, Reich D. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proc Natl Acad Sci U S A 2006;103(38):14068-14073.
2. Bonilla C, Panguluri RK, Taliaferro-Smith L, Argyropoulos G, Chen G, Adeyemo AA, Amoah A, Owusu S, Acheampong J, Agyenim-Boateng K, Eghan BA, Oli J, Okafor G, Abbiyesuku F, Johnson T, Rufus T, Fasanmade O, Chen Y, Collins FS, Dunston GM, Rotimi C, Kittles RA. Agouti-related protein promoter variant associated with leanness and decreased risk for diabetes in West Africans. International journal of obesity (2005) 2006;30(4):715-721.
3. Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, Oksenberg JR, Hauser SL, Smith MW, O'Brien SJ, Altshuler D, Daly MJ, Reich D. Methods for high-density admixture mapping of disease genes. American journal of human genetics 2004;74(5):979-1000.

APPENDICES: 2 articles attached

Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men

Matthew L. Freedman[a,b,c], Christopher A. Haiman[c,d], Nick Patterson[b,c], Gavin J. McDonald[b,e], Arti Tandon[b,e], Alicja Waliszewska[b,e,f], Kathryn Penney[b], Robert G. Steen[e,g], Kristin Ardlie[b,h], Esther M. John[i,j], Ingrid Oakley-Girvan[i,j], Alice S. Whittemore[j], Kathleen A. Cooney[k,l], Sue A. Ingles[d], David Altshuler[b,e,m,n], Brian E. Henderson[d], and David Reich[b,e,o]

[a]Department of Medical Oncology, Dana–Farber Cancer Institute, Boston, MA 02115; [b]Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142; [d]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089; Departments of [e]Genetics and [m]Medicine and [g]Biopolymers Facility, Harvard Medical School, Boston, MA 02115; [f]Laboratory of Molecular Immunology, Center for Neurologic Disease, Brigham and Women's Hospital, Boston, MA 02115; [h]Genomics Collaborative, Division of SeraCare Life Sciences, Inc., Cambridge, MA 02139; [i]Northern California Cancer Center, Fremont, CA 94538; [j]Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA 94305; [k]Departments of Medicine and Urology and [l]Comprehensive Cancer Center, University of Michigan, Ann Arbor, MI 48109; and [n]Center for Human Genetic Research and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114

Communicated by Eric S. Lander, Broad Institute, Cambridge, MA, July 12, 2006 (received for review May 24, 2006)

A whole-genome admixture scan in 1,597 African Americans identified a 3.8 Mb interval on chromosome 8q24 as significantly associated with susceptibility to prostate cancer [logarithm of odds (LOD) = 7.1]. The increased risk because of inheriting African ancestry is greater in men diagnosed before 72 years of age ($P < 0.00032$) and may contribute to the epidemiological observation that the higher risk for prostate cancer in African Americans is greatest in younger men (and attenuates with older age). The same region was recently identified through linkage analysis of prostate cancer, followed by fine-mapping. We strongly replicated this association ($P < 4.2 \times 10^{-9}$) but find that the previously described alleles do not explain more than a fraction of the admixture signal. Thus, admixture mapping indicates a major, still-unidentified risk gene for prostate cancer at 8q24, motivating intense work to find it.

association | human genetics

Prostate cancer is the most common noncutaneous malignancy among U.S. men, with an estimated 234,460 new cases and 27,350 deaths in 2006 (1). African Americans have the highest incidence of prostate cancer in the United States, ≈1.6-fold higher than European Americans (http://jncicancerspectrum.oxfordjournals.org/cgi/statContent/cspectfstat;18). The higher risk (2–4) prompted the hypothesis that genetic factors in part account for this difference. If there are genetic risk variants that differ substantially in frequency across populations, admixture mapping should have power to detect them.

The idea of admixture mapping is to screen through the genome of populations of mixed ancestry such as African Americans (5), searching for regions where the proportion of DNA inherited from either the ancestral European or African population is unusual compared with the genome-wide average. Admixture mapping requires a relatively small number of markers for a whole-genome scan: a couple of thousand, rather than the hundreds of thousands estimated to be necessary in nonadmixed populations (5, 6). Because the mixture between European and West-African populations occurred within the past 15 generations (5), stretches of DNA with contiguous European and African ancestry have not had much time to break up because of recombination and typically extend millions of base pairs. Admixture mapping therefore studies highly selected SNPs every few million base pairs (Mb), rather than every few thousand as with linkage disequilibrium mapping.

Although admixture mapping was first proposed >50 years ago (7) and has good power to detect risk variants that are strikingly different in frequency across populations (6, 8), it has not been practical until recently. Appropriate panels of markers (5), combined with analytical methods (8–10), made possible the first admixture scans (11, 12) in 2005. Here, we describe a whole-genome admixture scan focusing on prostate cancer, a disease that has long been considered a test case for admixture mapping because of its marked difference in incidence rates across populations. We identify a highly significant association at 8q24. The same broad region has recently been implicated in prostate cancer by Amundadottir *et al.* (13). In addition to providing independent evidence of a locus at 8q24, the present study provides two pieces of information. First, we show an association with earlier age of diagnosis. Second, we show that the alleles identified in the previous study are insufficient to explain more than a small fraction of the admixture signal. Thus, the causative alleles remain to be identified.

Results

We studied 1,597 prostate cancer cases and 873 controls, the majority of which were participants in the Multiethnic Cohort study (14) (810 cases and 730 controls) (Table 5, which is published as supporting information on the PNAS web site). The other samples came from six studies, including studies that specifically ascertained cases with high-grade tumors, advanced-stage disease, diagnosis at a young age, or occurrence in a family with multiple affected individuals (15–17) (Table 1). The present study was designed to include more cases than controls, because admixture mapping works by comparing the proportion of ancestry in cases to the rest of their own genomes. In principle, controls are not needed (6, 8); however, we included controls because they are useful for follow-up analyses (8).

All 2,470 samples (1,597 cases and 873 controls) were genotyped by using one of two panels of markers chosen to be highly different in frequency between West Africans and European Americans (5). A total of 1,792 samples were genotyped in the "phase 1" panel [previously used in a scan for multiple sclerosis genes (12)] and 1,266 SNPs passed quality filters and were used in analysis (Table 6, which is published as supporting information on the PNAS web site). The remaining 678 samples were typed in a second-generation "phase 2" panel that extracts more information per SNP; 1,365 SNPs passed quality filters and were used in analysis. The analysis combines information from both panels into a single logarithm of odds (LOD) score statistic at

Conflict of interest statement: No conflicts declared.

Abbreviations: LOD, logarithm of odds; OR, odds ratio.

[c]M.L.F., C.A.H., and N.P. contributed equally to this work.

[o]To whom correspondence should be addressed at: Department of Genetics, Harvard Medical School, New Research Building, 77 Avenue Louis Pasteur, Boston, MA 02115. E-mail: reich@receptor.med.harvard.edu.

Table 1. Characteristics of cases and controls from seven sources

Source	Location	Cases	Controls	Cases, % Euro. ± 1 SE	Controls, % Euro. ± 1 SE	Mean age diagnosis (range)	% with Gleason score >7	% with non-local tumors	% with prostate cancer in a first-degree relative	Decrease in peak LOD if these samples are removed*
Multiethnic Cohort	CA & HI	810	730	23.57 ± 0.50	25.42 ± 0.57	68 (46–85)	18	15	12	2.58
L.A. County Men's Health Study	CA	366	107	22.34 ± 0.83	26.37 ± 2.13	63 (42–88)	28	39	21	1.37
Study Early Onset Prostate Cancer	CA	104	—	20.89 ± 1.37	—	60 (45–65)	31	49	14	1.01
PCGP	MI	103	—	19.50 ± 1.01	—	55 (40–86)	11	29	39	1.15
Flint Men's Health Study	MI	85	—	18.05 ± 1.21	—	65 (47–77)	12	28	15	0.06
Bay Area Men's Health Study	CA	82	36	19.06 ± 1.52	20.13 ± 2.15	64 (44–78)	25	94	28	1.16
Genomics Collaborative	All U.S.	47	—	16.16 ± 1.51	—	62 (39–81)	14	38	28	0.57
Combined samples		1,597	873	22.11 ± 0.36	25.32 ± 0.55	65 (39–88)	21	29	18	7.14

PCGP, Prostate Cancer Genetics Project; Euro., European.

*To assess how much each of the seven cohorts contributes to the signal of association, we removed each from the main admixture scan (run no. 9 in Table 2) and assessed how the peak LOD score at 8q24 changes. All seven cohorts contribute positively.

each locus; observations >5 are considered strongly indicative of a disease locus (8). Formal significance is assessed by Bayesian methods. We take 10 to the power of the LOD score and average across points spaced every centimorgan across the genome. If the genome average is >100, then the Bayesian odds in favor of a disease locus is 100:1, and we interpret the data as showing significant evidence of a disease gene (8).

An initial admixture scan of 1,303 African-American prostate cancer cases produced a peak LOD score of 2.2 at 8q24. The signal was higher in a secondary analysis of individuals with a younger age at diagnosis, with the peak LOD score rising to 3.8 in the individuals who were <68 years of age (the threshold giving the strongest evidence of association). After genotyping 294 additional cases and 15 additional SNPs at 8q24 to obtain better local information about ancestry (see *Materials and Methods*), the peak LOD score increased to 4.1 in all cases and as high as 8.4 in the 1,176 who were diagnosed at <72 years of age. To correct for inflation of the score because of choosing the age threshold that gave the strongest significance, we integrated the evidence for association over an evenly spaced range of cutoffs (see *Materials and Methods* and Table 7, which is published as supporting information on the PNAS web site). This analysis yielded a peak LOD score of 7.1 (Fig. 1). Averaging 10 to the power of the LOD scores at equally spaced points genome-wide, we obtained a genome-wide average score of ≈19,000, exceeding the threshold of 100 for significance (8). After correcting for multiple hypothesis testing [by dividing by 4, because we tested four phenotypes (age, grade, stage, and familial disease) and focused on the one giving the strongest evidence], the odds in favor of a disease locus still greatly exceeded the threshold of 100 for significance.

The analysis in the previous paragraph used age of diagnosis as a covariate but did not directly test whether men with younger age of diagnosis have higher risk at 8q24 than older men. To formally test this hypothesis, we exploited the fact that ANCESTRYMAP software (see ref. 8 and http://genepath.med.harvard.edu/~reich) assigns scores for association to each individual separately (e.g., individual factors such as −0.02, 0.12) and then sums over all individuals to produce the total LOD score (Table 2). We rank-ordered the 1,588 cases in the scan for whom we had age information from youngest to oldest (Fig. 2). If the locus is not associated with age of diagnosis, the cumulative LOD should increase steadily to reach the total as additional samples are added. In fact, it rises to 5.4 LOD points above expectation at 71 years of age. To test whether this rise is significant, we permuted the data, reassigning ages of onset to different individuals (so that, in the randomized data, there could be no relationship between age of onset and allelic variation). In 1,000,000 permutations, only 318 showed a change in LOD score compared with the expectation exceeding the observed 5.4 ($P < 0.00032$). Repeating the analysis with a subset of samples obtained from a single prospective cohort [804 cases of African Americans with prostate cancer from the Multiethnic Cohort (MEC) Study], the association to age was also significant ($P < 0.0011$). These results indicate that there is a formally significant association of prostate cancer to ordering by age. We did not detect any associations when a similar analysis was applied to other subphenotypes: stage, grade, or family history (*Supplemental Note*

Table 2. Admixture scan summaries

	No. of cases	No. of controls	Peak LOD score
All prostate cancer cases	1,597	873	4.07
High grade (Gleason score >7)	316	873	2.68
Advanced stage (regional or metastatic cancer)	414	873	2.81
Family history (prostate cancer in a first-degree relative)	281	873	1.86
Age of diagnosis of <72 years and high-density genotyping at 8q24	1,176	873	8.39*
Drop out every even marker from run no. 5 to demonstrate independence of markers used	1,176	873	8.95*
Drop out every odd marker from run no. 5 to demonstrate independence of markers used	1,176	873	6.65*
Diagnosis at <72 years of age, high density and best model of 1.54-increased risk because of African ancestry	1,176	873	9.39*
Integrating over age-of-diagnosis cutoffs as a formal test for statistical significance	1,597	873	7.14*

*Indicates a scan that meets formal criteria for genome-wide statistical significance.

GENETICS

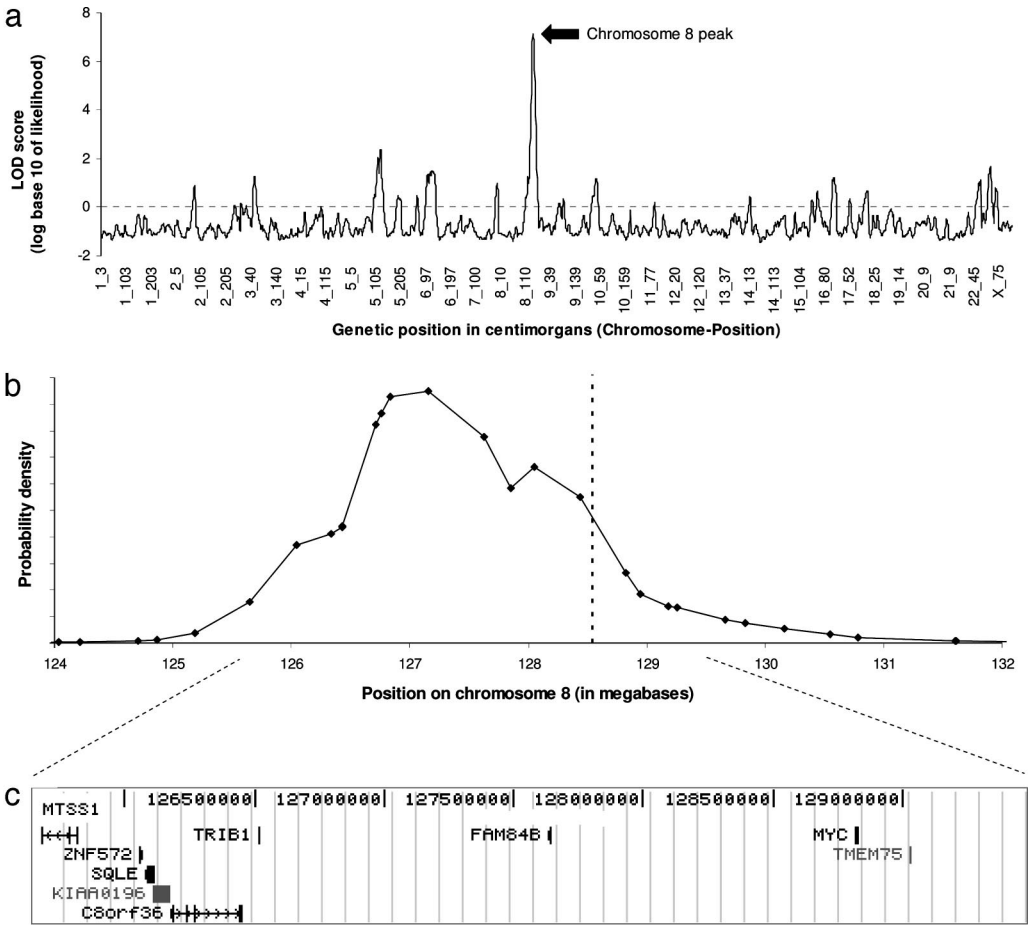

Fig. 1. Summary of results for the whole-genome admixture scan and characteristics of the 8q24 peak of association. (*a*) We present the LOD score at equally spaced points across the genome. The chromosome 8 peak is marked by a rise to 7.14. (*b*) We can use the data to calculate a probability distribution for the position of the peak. It aligns with the microsatellite and SNP recently associated with prostate cancer by Amundadottir *et al*. (13) (dashed line). (*c*) The 95% credible interval spans 3.8 Mb (125.68–129.48 Mb in build 35 of the human reference sequence) and contains nine known genes, including the c-MYC oncogene (diagram taken from http://genome.ucsc.edu) (data from the May 2004 genome assembly).

1 in *Supporting Text*, which is published as supporting information on the PNAS web site).

To explore how much of the increased incidence of prostate cancer in African-American men might be explained by African (as compared with European) ancestry at 8q24, we evaluated the risk for individuals carrying zero, one, and two chromosomes with African ancestry at the locus. Each African-derived chromosome is associated with ≈1.54-fold increased risk in younger individuals (90% credible interval 1.38–1.74) (*Supplemental Note 2*). We also estimated the proportion of control samples with zero, one, and two African-derived chromosomes, respectively (6.4%, 37.8%, and 55.8%, respectively). Extrapolating to the broader African-American population, the prostate cancer incidence in all African Americans I_{ALL} is higher than the incidence I_{EE} in individuals who inherited two European-derived chromosomes at the locus by a factor of $[(0.064)\ (1) + (0.378)(1.54) + 0.558(1.54^2)] = 1.969$. Thus, the fraction of all prostate cancer incidence for African Americans <72 years of age that could be explained by ancestry at this locus is $(I_{ALL} - I_{EE})/I_{ALL} = 1 - (1/1.969) = 49\%$ (with a 90% credible interval of 39–59%). Thus, if it were possible to develop a treatment that reduced prostate cancer risk in the African-American population to the level that is seen in men who carry two copies of 8q24 inherited from recent European ancestors, the rate of prostate cancer would decrease by ≈49%. The total risk for prostate cancer that can be attributed to 8q24 in African-American men

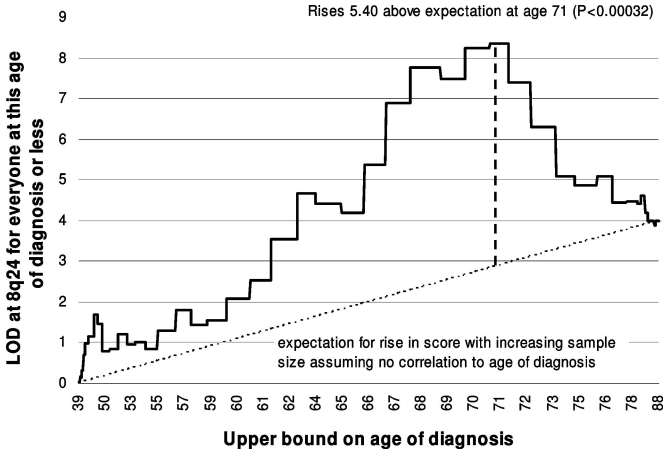

Fig. 2. To formally test for a relationship between age of onset and contribution to the chromosome 8 locus, we rank-ordered the individuals by age of onset and then calculated a score for increasing age cutoffs. The score rises to 5.40 above the expectation for 1,176 individuals diagnosed at <72 years of age. To evaluate whether this rise is unexpected, we permuted the data 1,000,000 times, randomizing scores with respect to individuals' ages of onset (guaranteeing that there is no relationship between age of diagnosis and contribution to the evidence of association). In only 318 of 1,000,000 permutations did we see a rise as high as in our data ($P < 0.00032$).

<72 years of age is still greater, because alleles at 8q24 increase prostate cancer risk even in chromosomes of entirely European origin (13). Thus, 8q24 has a major effect on population risk of prostate cancer, especially in younger African Americans.

Using the LOD scores at 8q24, we also calculated a posterior probability distribution to estimate the position of the disease-causing variants (see Fig. 1*b* and *Materials and Methods*). The 95% credible interval spans 3.80 Mb, from 125.68–129.48 Mb in build 35 of the human genome reference sequence (13.9 cM) and contains nine known genes (Fig. 1*c*). However, the admixture scan does not provide information about which gene or alleles within the locus confer risk.

Independently of this study, Amundadottir *et al*. (13) reported an SNP allele [A at rs1447295; odds ratio (OR) = 1.51; $P < 1.0 \times 10^{-11}$] and a microsatellite allele (−8 at DG8S737; OR = 1.62; $P < 2.7 \times 10^{-11}$) that map to the same region as the admixture peak (at 128.546 Mb and 128.554 Mb, respectively) and are highly associated with prostate cancer. The effect of the −8 allele was observed in European and African Americans, whereas the A allele effect was detected only in European-derived populations. The authors did not show, however, that either allele was causally involved in disease but instead suggested that they were both in linkage disequilibrium with an as-yet-unidentified causal variant. They also did not identify which gene in the region might be responsible for prostate cancer risk.

To directly compare the results of the two studies, we tested the previously associated alleles in the African-American cases and controls [excluding samples from Michigan, because they overlap those studied by Amundadottir *et al*. (13); see *Materials and Methods*]. The goal was to test whether the −8 allele at DG8S737 contributes to disease risk in African Americans beyond the risk that can be accounted for by the admixture signal. (*Supplemental Note 3*). We were concerned that the previously detected association in African Americans by Amundadottir *et al*. (13) ($P < 0.0022$, estimated OR of 1.60) might simply reflect an admixture signal across a large region (because of systematic differences in ancestry between cases and controls across several million base pairs of 8q24) and thus might not provide fine-mapping information in African Americans. Although Amundadottir *et al*. (13) tested for mismatching of cases and controls in overall proportion of ancestry, they did not control for a local rise in African ancestry throughout 8q24 in cases but not controls. Such a rise would be expected to cause thousands of alleles in the region that just happen to be more frequent in African Americans (including the microsatellite −8 allele) to show association with prostate cancer. When we correct for this effect in the African-American samples from the present study (*Supplemental Note 3*), we find that the contribution of the −8 allele to risk is nonsignificant ($P = 0.22$) (Table 3). The OR of 0.93–1.17 (95% credible interval) also rules out the OR = 1.60 reported in African-Americans (13).

We next expanded the replication analysis to the four ethnicities in the MEC other than African Americans, by genotyping rs1447295 in 1,614 prostate cancer cases and 1,547 controls from these populations. The evidence for association is significant overall ($P < 4.2 \times 10^{-9}$), as well as separately in each group: Japanese Americans ($P < 0.00034$), Native Hawaiians ($P < 0.00015$), Latino Americans ($P < 0.0014$), and European Americans ($P < 0.022$) (Table 4). This analysis replicates the association identified by Amundadottir *et al*. (13), although we did not test for the possible confounding factor of population stratification. Interestingly, we do not replicate the association to tumor grade (Gleason ≥8 vs. Gleason <8; $P = 0.47$) reported by Amundadottir *et al*. (13).

These results confirm the finding of Amundadottir *et al*. (13) that the 8q24 locus is important in prostate cancer. However, the alleles they reported do not explain the admixture signal (*Supplemental Note 4* and Table 3). The specific variants causing increased risk for prostate cancer in African American because of 8q24 thus remain to be identified.

Table 3. Allelic association tests in African Americans adjusting for local rise in African ancestry

	Cases	Controls	*P* value	OR (95% CI)
A allele at rs1447295	989	804	0.15	1.05 (0.95–1.16)
−8 allele at DG8S737	966	797	0.22	1.05 (0.93–1.17)
Haplotype of A and −8*	902	776	0.31	1.03 (0.90–1.18)

Cases diagnosed at <72 years of age and all controls. *P* values are one-tailed, testing for the previously associated allele (14) being more common (*Supplemental Note 3*). CI, confidence interval.

*For the haplotype test, we phased the cases and controls together, before carrying out the association analysis.

Discussion

We have used admixture mapping to identify a locus at 8q24 that substantially affects risk for prostate cancer. We highlight four findings.

First, this study shows that admixture mapping can be a powerful and practical way to map genetic variants for complex disease (5, 18). The results motivate the application of admixture mapping to other disorders, especially those like prostate cancer in which incidence varies across populations. These results also highlight the scientific value of studies to find disease genes in specific ethnic groups, such as African Americans.

Second, we show that the 8q24 locus contributes to a major increased risk for prostate cancer in African Americans with African ancestry at 8q24. The difference between these individuals and African Americans with European ancestry at 8q24 explains a large proportion of prostate cancer in younger African Americans. If one could intervene medically to reduce the risk for prostate cancer in African Americans <72 years of age to what would be expected if all African Americans had European ancestry at the locus, the incidence in men <72 years of age would decrease by approximately 49%. We also show that the admixture signal at 8q24 cannot be explained by the alleles identified by Amundadottir *et al*. (13); instead, there must be major, unmapped risk alleles at the locus.

Third, we detect a highly significant association of 8q24 with age. This finding is intriguing because it is known epidemiologically that the differential incidence of prostate cancer in African versus European Americans is greater at younger ages and is attenuated with older age (ref. 19; http://jncicancerspectrum.oxfordjournals.org/cgi/statContent/cspectfstat;18). Surveillance, Epidemiology, and End Results (SEER) Program registry data indicate that, for men diagnosed at <55 years of age, African Americans have a 2.27-fold higher rate than European Americans, but the ratio decreases to 1.48-fold for men diagnosed at ≥75 years of age (19). Genetic variation at 8q24 may be responsible for part of this effect.

Fourth, we identify a 3.8-Mb interval containing nine known genes that is likely to harbor variant(s) explaining the admixture peak. This is a tractable region for follow-up analysis. Somatic genetic data independently highlight the 8q24 region as one of the most frequently amplified regions in prostate cancer tumors (20, 21). The *c-MYC* oncogene, a key regulator in cellular proliferation, lies within the peak. Overexpression of *c-MYC* has been shown to induce tumors in mice and to create a cancer phenotype in benign prostatic epithelium (22, 23). It is possible that *c-MYC* could be the gene responsible for the prostate cancer risk, but no structural or regulatory variant has yet been identified.

Follow-up work will be necessary to identify the as-yet-undiscovered causal risk variant(s) at 8q24. Ultimately, discovering the causal gene(s) at 8q24 may translate into better

Table 4. rs1447295 association in the Multiethnic Cohort

Group within the Multiethnic Cohort	No. of samples		Frequency of A allele, %		*P* value (one-tailed)	OR (95% confidence interval)
	Cases	Controls	Cases	Controls		
Native Hawaiians	70	68	37.0	16.2	0.00015	3.02 (1.66–5.50)*
Japanese Americans	449	465	23.8	17.2	0.00034	1.48 (1.18–1.86)*
Latino Americans	640	567	13.5	9.5	0.0014	1.48 (1.14–1.91)*
European Americans	455	447	13.1	10.0	0.022	1.35 (1.01–1.80)*
All samples together	1,614	1,547			4.2×10^{-9}	1.36 (1.22–1.51)†

*OR estimated by using logistic regression adjusted for age.
†OR estimated by using logistic regression adjusted for age as well as ethnicity.

understanding of prostate cancer and may play a role in strategies for screening of the population and identifying new targets for treatment and prevention.

Materials and Methods

Samples. Samples were derived from seven sources (Table 1). The largest number came from the Multiethnic Cohort (MEC), a prospective cohort that began in 1993 and is still ongoing, which ascertains prostate cancer cases and controls by linking to databases from the California Cancer Registry, the Los Angeles County Cancer Surveillance Program, and the Hawaii Cancer Registry (14). The samples used in the admixture scan were all African-American cases and controls; however, for the validation genotyping of the rs1447295 SNP, we also genotyped prostate cancer cases and controls from four other ethnicities in the MEC: European Americans, Latino Americans, Japanese Americans, and Native Hawaiians. The second largest number of samples came from the Los Angeles County Men's Health Study (1999–2002), which was enriched for individuals with advanced-stage or high-grade prostate cancer, as identified through hospitals and private histopathology laboratories in Los Angeles County. The Bay Area Men's Health Study (15) (1997–2000) was enriched for individuals with regional- or distant-stage disease. The Study of Early Onset Prostate Cancer (1993–1995) was based in the San Francisco–Oakland Bay Area and included only individuals with histologically confirmed prostate cancer who were <66 years of age at diagnosis. The Genomics Collaborative, Ltd. samples were obtained from consenting individuals undergoing surgery for prostate cancer throughout the U.S. and were provided to this study at no cost by means of an academic collaboration. The Flint Men's Health Study samples (1996–2002) were obtained through a case-control study of prostate cancer in Genesee County, Michigan. The University of Michigan Prostate Cancer Genetics Project (PCGP) samples were obtained from an ongoing family-based study of prostate cancer susceptibility. PCGP cases have a family history of prostate cancer or early age at diagnosis defined as <55 years of age (we analyzed data only from the man with the youngest age of diagnosis in each family). We note that both the Flint Men's Health Study and PCGP samples (16, 17) overlap with those studied by Amundadottir *et al*. (13). The samples were provided by K.A.C. for replication purposes blinded to the locus under study. The results reported here, which also use a different type of information to localize disease genes (admixture linkage disequilibrium), are thus fully independent.

Genotyping. The phase 1 and phase 2 panels of SNPs were both genotyped by using the Illumina BeadLab genotyping platform (24) [supplemented for phase 1 by Sequenom MassARRAY genotyping (25)]. At the 8q24 peak, we genotyped an additional 15 SNPs using Sequenom technology to extract maximal information about ancestry [these SNPs were chosen to have high frequency differentiation between the European and West-African populations (5) based on data from the Human Haplotype Map (26)]. We used previously described protocols to remove SNPs that did not perform well in genotyping, that were in linkage disequilibrium with each other in the ancestral European and West-African populations, or that did not seem to have appropriate intermediate frequencies in the African Americans compared with the ancestral populations (12). The rs1447295 genotyping was carried out by using the Applied Biosystems Inc. (ABI, Foster City, CA) Assay-on-Demand technology following the manufacturer's recommended protocol, and all of the African Americans were also genotyped at rs1447295 by using Sequenom technology. The DG8S737 genotyping was carried out by using ABI True Allele PCR Premix, with 5-pmol forward (5′-6FAM-TGATGCACCACAGAAACCTG-3′) and 5-pmol reverse (5′-GTTTCAAGGATGCAGCTCACAACA-3′) primers, and 60 ng of DNA per reaction. Reactions were analyzed on an ABI3730xl DNA Analyzer. Samples were scored by the ABI GeneMapper V3.7 software, with all genotypes confirmed by an experienced technician. To check the microsatellite genotyping results, we compared 168 samples that overlapped between this study and that of Amundadottir *et al*. (13) (data provided by K.A.C.); only five comparisons were inconsistent.

Admixture Analysis. We used the ANCESTRYMAP software (8) to carry out the screens for association with prostate cancer. ANCESTRYMAP calculates a statistic for association at every position in the genome, under a prespecified family of risk models, calculating the likelihood of the data at the locus under an average of disease models versus the likelihood of the data if the locus has nothing to do with disease (the log base 10 of this is the LOD score). For most runs, we assume equally likely models of 0.3-, 0.4-, 0.6-, 0.7-, 0.8-, 1.2-, 1.5-, and 2-fold increased risk because of each copy of a European allele. This family of models reflects the hypothesis that African-derived alleles are more likely to confer risk but also tests for the alternative possibility. To obtain an overall assessment of the evidence for a disease locus anywhere in the genome, we average the factors for association at each point separately, providing a genome-wide assessment of whether there is a locus in the genome affecting risk.

Admixture Scan Accounting for Age of Diagnosis. We carried out an admixture scan taking into account the possibility that individuals with a younger age of diagnosis contribute a more powerful admixture signal, while not inappropriately inflating the signal of association by picking the cutoff giving the strongest signal. We ran 22 independent scans for all individuals in the data set with diagnosis at <50, <53, <56, <57, <59, <60, <61, <62, <63, <64, <65, <66, <67, <69, <70, <71, <73, <74, <75, <76, and <78 years of age, as well as all cases (Table 7). Approximately 73 new samples were added in for each consecutive run. We then averaged

the genome scores for association, which gives a statistically appropriate assessment of the evidence for association.

Permutation Analysis to Test Whether Some Phenotypes Contribute Unduly to the Signal of Association at 8q24. To test whether the correlation of the 8q24 admixture association with a phenotype is significant, we carried out permutation analyses, considering separately the effect of stage of disease, grade of tumor, family history, and age of diagnosis (Fig. 2 and *Supplemental Note 1*). For each phenotype, we rank-ordered individuals by their values of the phenotype. We then calculated a cumulative LOD score at SNP rs780321 (used to mark the peak) for all individuals below each cutoff. We recorded the greatest excess or shortfall of the cumulative LOD score compared with the expectation if it increased linearly. We then wrote a PERL script to randomly permute the values of the phenotype over the samples, eliminating any relationship between the phenotype and score. A P value was calculated as the fraction of 1,000,000 permutations that produced a score for association as extreme as the data.

Inferring the Position of the Disease Locus. To infer the position of the disease locus, we note that the LOD scores at each point of the genome can be taken to the power of 10 to give the relative probability of that locus containing the disease allele. After normalization, this calculation provides a probability distribution for the position of the locus. A 95% credible interval is obtained from the central area under the peak (Fig. 1*c*).

We thank the men with and without prostate cancer who participated in this study, Eric Lander and two reviewers for comments and criticism, Loreall Pooler and David Wong from the University of Southern California Genomics Laboratory for help with sample handling and genotyping, Courtney Montague at Harvard Medical School for assistance with genotyping, and the National Center for Research Resources Center for Genotyping and Analysis at the Broad Institute, without which this work would not have been possible. The genotyping for this work was supported by National Institutes of Health (NIH) Grant CA63464 (to B.E.H., C.A.H., D.A., and D.R.). M.L.F. was supported by a Department of Defense Health Disparity Training-Prostate Scholar Award (DAMD 17-02-1-0246), by a Howard Hughes Medical Institute physician postdoctoral fellowship, and by Dana–Farber/Harvard Partners Cancer Care Prostate Specialized Programs of Research Excellence (SPORE). N.P. was supported by NIH Career Transition Award HG02758. E.M.J. and S.A.I. were supported by California Cancer Research Program Grants 99-00527V-10182 and 99-00524V-10258, respectively. The Flint Men's Health Study was supported by the University of Michigan SPORE in Prostate Cancer (CA69568), the University of Michigan Department of Urology, and the University of Michigan Comprehensive Cancer Center. K.A.C. was supported by NIH Awards CA69568 and CA79596, and I.O.-G. and A.S.W. were supported by NIH Award CA67044. D.A. is a Charles E. Culpeper Scholar of the Rockefeller Brothers Fund and a Burroughs Wellcome Fund Clinical Scholar in Translational Research. D.R. is the recipient of a Burroughs Wellcome Career Development Award in the Biomedical Sciences.

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, Thun MJ (2006) *CA Cancer J Clin* 56:106–130.
2. Bunker CH, Patrick AL, Konety BR, Dhir R, Brufsky AM, Vivas CA, Becich MJ, Trump DL, Kuller LH (2002) *Cancer Epidemiol Biomarkers Prev* 11:726–729.
3. Freedland SJ, Isaacs WB (2005) *Prostate* 62:243–252.
4. Simard J, Dumont M, Labuda D, Sinnett D, Meloche C, El-Alfy M, Berger L, Lees E, Labrie F, Tavtigian SV (2003) *Endocr Relat Cancer* 10:225–259.
5. Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, Kessing BD, Malasky MJ, Scafe C, De Jager PL, *et al.* (2004) *Am J Hum Genet* 74:1001–1013.
6. McKeigue PM (1997) *Am J Hum Genet* 60:188–196.
7. Rife DC (1954) *Am J Hum Genet* 6:26–33.
8. Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, Oksenberg JR, Hauser SL, Smith MW, O'Brien SJ, Altshuler D, *et al.* (2004) *Am J Hum Genet* 74:979–1000.
9. Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM (2004) *Am J Hum Genet* 74:965–978.
10. Montana G, Pritchard JK (2004) *Am J Hum Genet* 75:771–789.
11. Zhu X, Luke A, Cooper RS, Quertermous T, Hanis C, Mosley T, Gu CC, Tang H, Rao DC, Risch N, *et al.* (2005) *Nat Genet* 37:177–181.
12. Reich D, Patterson N, De Jager PL, McDonald GJ, Waliszewska A, Tandon A, Lincoln RR, DeLoa C, Fruhan SA, Cabre P, *et al.* (2005) *Nat Genet* 37:1113–1118.
13. Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Cazier JB, Sainz J, *et al.* (2006) *Nat Genet* 38:652–658.
14. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, Stram DO, Monroe KR, Earle ME, Nagamine FS (2000) *Am J Epidemiol* 151:346–357.
15. John EM, Schwartz GG, Koo J, Van Den Berg D, Ingles S (2005) *Cancer Res* 65:5470–5479.
16. Cooney KA, Strawderman MS, Wojno KJ, Doerr KM, Taylor A, Alcser KH, Heeringa SG, Taylor JM, Wei JT, Montie JE, *et al.* (2001) *Urology* 57: 91–96.
17. Cooney KA, McCarthy JD, Lange E, Huang L, Miesfeldt S, Montie JE, Oesterling JE, Sandler HM, Lange K (1997) *J Natl Cancer Inst* 89:955–959.
18. Kittles RA, Chen W, Panguluri RK, Ahaghotu C, Jackson A, Adebamowo CA, Griffin R, Williams T, Ukoli F, Adams-Campbell L, *et al.* (2002) *Hum Genet* 110:553–560.
19. Whittemore AS, Keller JB, Betensky R (1991) *J Natl Cancer Inst* 83:1231–1235.
20. van Duin M, van Marion R, Vissers K, Watson JE, van Weerden WM, Schroder FH, Hop WC, van der Kwast TH, Collins C, van Dekken H (2005) *Genes Chromosomes Cancer* 44:438–449.
21. Visakorpi T, Kallioniemi AH, Syvanen AC, Hyytinen ER, Karhu R, Tammela T, Isola JJ, Kallioniemi OP (1995) *Cancer Res* 55:342–347.
22. Ellwood-Yen K, Graeber TG, Wongvipat J, Iruela-Arispe ML, Zhang J, Matusik R, Thomas GV, Sawyers CL (2003) *Cancer Cell* 4:223–238.
23. Williams K, Fernandez S, Stien X, Ishii K, Love HD, Lau YF, Roberts RL, Hayward SW (2005) *Prostate* 63:369–384.
24. Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, Hansen M, Steemers F, Butler SL, Deloukas P, *et al.* (2003) *Cold Spring Harbor Symp Quant Biol* 68:69–78.
25. Tang K, Fu DJ, Julien D, Braun A, Cantor CR, Koster H (1999) *Proc Natl Acad Sci USA* 96:10016–10020.
26. The International HapMap Consortium (2005) *Nature* 437:1299–1320.

E-Cadherin Polymorphisms and Haplotypes Influence Risk for Prostate Cancer

Carolina Bonilla,[1,2] Tshela Mason,[2] Layron Long,[2] Chiledum Ahaghotu,[2,3] Weidong Chen,[2] Aiqiu Zhao,[2] Aoua Coulibaly,[2] Frankly Bennett,[4] William Aiken,[4] Trevor Tullock,[4] Kathleen Coard,[4] Vincent Freeman,[5,6] and Rick A. Kittles[1,2,3]*

[1]*Human Cancer Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio*
[2]*National Human Genome Center, College of Medicine, Howard University, Washington, DC*
[3]*Division of Urology, Howard University Hospital, Washington, DC*
[4]*Tropical Metabolism Research Institute, University of the West Indies, Mona, Kingston, Jamaica, West Indies*
[5]*Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, Illinois*
[6]*Department of Urology, Loyola University Stritch School of Medicine, Maywood, Illinois*

BACKGROUND. The E-cadherin (*CDH1*) gene has been implicated in prostate cancer (PCA) risk, however, the exact mechanism is unknown. Several polymorphisms, such as the C/A variant −160 base pairs from the transcription start site, in the *CDH1* gene promoter region have been associated with cancer risk, mainly in European descent populations.

METHODS. We screened the entire coding region and 3.0 kilobases of the *CDH1* promoter for polymorphisms in 48 African Americans using dHPLC. Twenty-one (21) polymorphisms were observed. Four polymorphisms, including −160C/A, were genotyped in a genetic association study using incident PCA cases (N = 427) and unaffected controls (N = 337) of similar age from three different ethnic groups consisting of African Americans, Jamaicans, and European Americans.

RESULTS. We observed a significantly higher frequency of the −160A allele among European American PCA patients (27.5%) compared to the control group (19.7%) ($P = 0.04$). More importantly, among men of European ancestry under the age of 65 who possess the −160 A allele there was over three times increased risk for prostate cancer ($P = 0.05$). Also, the AACT haplotype bearing the −160A allele was significantly associated with PCA in European Americans ($P = 0.04$).

CONCLUSIONS. Our data indicate that *CDH1* likely is a low-penetrant PCA susceptibility gene, however, population differences in linkage disequilibrium within the *CDH1* gene region may influence the effect of susceptibility alleles such as −160A. *Prostate 66: 546–556, 2006.*

KEY WORDS: **CDH1; prostate cancer; African Americans; single nucleotide polymorphisms (SNP); tumor suppressor; haplotypes**

INTRODUCTION

Prostate cancer (PCA) is one of the most common malignancies among men in developed countries [1]. Risk factors for PCA development include advanced age, ethnicity, and a positive family history. To date, a few susceptibility genes have been identified, although no major predisposition locus has been observed so far [2]. Thus, it has been suggested that low-penetrance susceptibility genes with higher

Grant sponsor: NIH; Grant numbers: RR03048-13S1, 1U54CA91431-01; Grant sponsor: Department of Defense (to RAK); Grant numbers: DAMD17-00-1-0025, DAMD 17-02-1-0067; Grant sponsor: Department of Defense (to VF); Grant number: DAMD17-00-1-0029.

*Correspondence to: Rick A. Kittles, PhD, Human Cancer Genetics, Comprehensive Cancer Center, The Ohio State University, 494 Tzagournis Medical Research Facility, 420 W. 12th Avenue, Columbus, OH 43210. E-mail: rick.kittles@osumc.edu
Received 28 August 2005; Accepted 27 September 2005
DOI 10.1002/pros.20374
Published online 21 December 2005 in Wiley InterScience (www.interscience.wiley.com).

population frequencies may be relevant in the determination of PCA risk in combination with environmental factors [3]. Several common, low penetrant genes have been identified as potential PCA susceptibility genes. These candidate genes include *SRD5A2* (MIM 607306), *CYP3A4* (MIM 124010), *CYP3A5* (MIM 605325), *VDR* (MIM 601769), and *E-cadherin* (MIM 192090).

E-cadherin (CDH1) is an adhesion glycoprotein found in epithelial tissues where it promotes cell–cell unions known as adherens junctions [4]. As malignant cells show poor adhesion properties in addition to loss of differentiated epithelial morphology and increased cellular motility, it has been proposed that CDH1 may play a role in tumor initiation and progression [4]. In fact, loss of CDH1 expression is believed to be the fundamental step in the disruption of tight intercellular contact that leads to the invasive and metastatic state of tumors [5]. Earlier studies have provided evidence for a role of CDH1 as a tumor suppressor in several human cancers, where loss or reduced expression of CDH1 has been demonstrated [6–8]. Among the mechanisms responsible for aberrant CDH1 expression are loss of heterozygosity, promoter hypermethylation, and somatic and germline mutations. Germline mutations in *CDH1* are present in about one-third of families affected of hereditary diffuse gastric cancer syndrome (HDGC) [9], and have been shown to affect cell motility and invasion [10]. In lobular breast cancer, complete loss of E-cadherin expression is a characteristic of about 80% of the carcinomas, where *CDH1* mutations are fairly frequent [11]. In turn, hypermethylation of CpG islands in the *CDH1* promoter leading to decreased gene expression has been found in bladder, breast, colon, liver, oral, and prostate cancers [12]. Furthermore, analysis of methylation levels in the *CDH1* promoter of PCA cell lines has shown that methylation of *CDH1* correlates with tumor progression [13].

The *CDH1*gene is located at 16q22.1 and consists of 16 exons spanning approximately 100 kb of genomic DNA. Several polymorphisms, germline and somatic mutations have been identified within its coding regions [9,14,15]. A −160C/A polymorphism in the promoter was described by Li et al. [13], who reported reduced transcriptional activity of ~70% of the A allele compared to the C allele. Further research revealed that the A allele conferred a modest increased risk of PCA in European populations [16–18]. Yet, no association between −160C/A genotype and tumor progression or metastasis was evident among Slovenians [18]. In Japanese, conflicting results were reported [19,20]. However, no studies have explored the relationship of *CDH1* variants with susceptibility to cancer in populations of African descent. Given that prostate cancer exhibits a higher prevalence in African descent populations compared to the European American population, we were interested in establishing whether sequence variants in *CDH1* could explain in part the elevated risk.

The objective of this study was to screen the *CDH1* gene for additional sequence polymorphisms in men of African descent and test for *CDH1* effects on PCA risk by performing case-control association analyses in African Americans, European Americans, and Jamaicans of African descent. Our results reveal modest effects of the −160C/A allele on PCA risk among European Americans. In addition it is likely that *CDH1* alleles vary in their effect on PCA risk due to the influence of other linked functional polymorphisms and possibly other genes (gene–gene interactions).

MATERIALS AND METHODS

Study Populations

Unrelated men were enrolled from three sites for genetic association studies of risk factors for PCA. All PCA cases were between 40 and 85 years of age and were diagnosed with PCA within a year prior to recruitment. The first group of men consisted of 231 African Americans (119 PCA patients and 112 male controls) recruited from the Washington, DC area through the Division of Urology at the Howard University Hospital and/or PCA screening at the Howard University Cancer Center. Unaffected African American male volunteers were enrolled among individuals undergoing regular physical exams at the Division of Urology at Howard University Hospital and/or men participating in screening programs for PCA at the Howard University Cancer Center. The screening program was demographically similar to the patient population seen in the Division of Urology clinics. The recruitment of controls occurred concurrently with individuals recruited with PCA. Mean age of the African American PCA patients was 65.1 ± 0.9 and among controls 67.2 ± 1.1.

The second group consisted of PCA cases and age and ethnicity matched controls (89 PCA patients and 123 unaffected male controls). All men in this group were of African descent and resided on the Caribbean island of Jamaica. The Jamaican men were recruited from the University Hospital of the West Indies in Kingston, Jamaica during the year 2000. Each case subject was diagnosed with prostate carcinoma by a pathologist. Men free of prostate cancer were also recruited from prostate cancer screening programs on the island. Mean age of the Jamaican PCA patients was 67.1 ± 1.4 and among controls 65.4 ± 1.0.

The third group consisted of 321 European American men (219 PCA patients and 102 unaffected male controls) recruited from the Chicago metropolitan area

through a department of urology at a single academic institution (Loyola University Medical Center). All cases were men diagnosed with clinically localized PCA and awaiting radical prostatectomy. Clinically evaluated healthy male controls of European ancestry were also recruited from the Chicago metropolitan area. Mean age of the PCA patients was 61.0 ± 0.6 and among controls 63.9 ± 1.0.

Blood samples were collected from each subject. Clinical characteristics including Gleason grade, prostate specific antigen (PSA), and age at diagnosis were obtained from medical records. Disease aggressiveness was defined as ''low'' (Gleason grade <7) or ''high'' (Gleason grade ≥ 7). All controls had PSA levels <4.0 ng/ml and normal digital rectal exams (DRE) (Table I). Individuals diagnosed with benign prostatic hyperplasia (BPH) were not considered in this analysis. Howard University Institutional Review Board approved the study and written consent was obtained from all subjects.

SNP Discovery

Genomic DNA was isolated from lymphocytes using standard proteinase K digestion, cell lysis, protein precipitation, and DNA precipitation. A total of 3,000 base pairs upstream of the ATG start site and all 16 exons of *CDH1* were screened for DNA sequence variation by denaturing high-performance liquid chromatography (dHPLC) using the Wave™ DNA fragment analysis system (Transgenomic, Omaha, NE) according to the manufacturer's instructions. Information about the primers and PCR conditions used to amplify *CDH1* promoter and exon fragments are available upon request.

Genomic DNA from 48 African American individuals (24 cases and 24 controls) was used for SNP detection, and four SNPs were selected to be typed in the combined population set. These variants included two within the promoter region (-1004A/T and -160C/A) and two within intron 1 (IVS1 + 5C/G and IVS1 + 6T/C).

Genotyping

CDH1 -1004 SNP was genotyped by direct sequencing using an ABI 377 DNA sequencer (PE Biosystems). The promoter SNP -160C/A was genotyped by restriction endonuclease digestion and conventional agarose gel electrophoresis. Polymerase chain reaction (PCR) primers and protocol, as well as digestion conditions, were previously reported by Verhage et al. [16].

Genotyping of *CDH1* variants IVS1 + 5C/G and IVS1 + 6T/C was performed using Pyrosequencing™ (Pyrosequencing, AB, Uppsala, Sweden) according to standard protocols with the PSQ96 automated

TABLE I. Characteristics of Study Populations

Population trait	Cases	Controls
All populations		
Number of subjects	427	337
Mean age in years ± SE[a]	63.5 ± 0.5	66.0 ± 0.6
Mean serum PSA in ng/ml ± SE[b]	94.9 ± 21.0	1.1 ± 0.1
Gleason grade ≥7 (%)	185 (58)	—
African Americans		
Number of subjects	119	112
Mean age in years ± SE	65.1 ± 0.9	67.2 ± 1.1
Mean serum PSA in ng/ml ± SE[b]	55.1 ± 16.2	1.1 ± 0.1
Gleason grade ≥7 (%)	28 (44)	—
European Americans		
Number of subjects	219	102
Mean age in years ± SE[a]	61.0 ± 0.6	63.9 ± 1.0
Mean serum PSA in ng/ml ± SE[b]	20.8 ± 8.9	1.2 ± 0.8
Gleason grade ≥7 (%)	105 (61)	—
Jamaicans		
Number of subjects	89	123
Mean age in years ± SE	67.1 ± 1.2	65.4 ± 1.0
Mean serum PSA in ng/ml ± SE[b]	123.1 ± 37.2	1.0 ± 0.4
Gleason grade ≥7 (%)	52 (65)	—

[a]Mean age difference between cases and controls significant at $P \leq 0.05$.
[b]Serum PSA measured at time of diagnosis for cases and at most recent clinical visit for controls.

Pyrosequencing instrument [21]. Fragments were amplified using the following forward and reverse primers: *CDH1*-P1F 5′-AGA CTC CAG CCC GCT CCA-3′ and *CDH1*-P1R 5′-biotin- GGC CCG AAT GCG TCC CT-3′. The following pyrosequencing primer was used to genotype both intronic variants: *CDH1*-P1pyro 5′-CTG CTG CTG CAG GTA-3′. All samples were genotyped twice directly from genomic DNA. Control DNAs of known genotype were also included. The control genotypes were confirmed by direct DNA sequencing. Genotypes from the repeat assays were 100% concordant with initial genotypes. Presumed alterations in the transcription factor binding sites due to each variant were predicted using the gene regulation website (http://www.gene-regulation.com) [22].

Statistical Analysis

Genotype and allele frequencies were calculated for each population. Frequency differences between populations and between patients and controls within populations were examined by contingency table analysis. Hardy–Weinberg equilibrium was evaluated in each group using a Chi-square test.

Binomial logistic regression was performed using SPSS software (v.11) to test for association of *CDH1* genotypes with disease status. We analyzed all populations combined, with adjustment for ethnicity and age, and each population separately controlling for age. In addition, we used SAS/Genetics (SAS Institute, Cary, NC) to perform an allele and genotype case-control test and a linear trend test. Specifically, the genotype test evaluates dominant effects of alleles, whereas the allele and the linear trend tests assess additive allelic effects.

Haplotype reconstruction and assessment of haplotype frequency differences between cases and controls, as well as pairwise linkage disequilibrium for each pair of SNPs, were obtained using SAS/Genetics, which implements an EM algorithm [23–25]. Two-sided Pearson Chi-square, odds ratios, and *P*-values were determined for the most frequent haplotypes (>5%) from comparisons between cases and controls using all populations combined and each particular population.

Additional analyses involved testing for association of *CDH1* polymorphisms and PCA susceptibility after stratification of the sample by age (<60/≥60 and <65/≥65 years), as well as evaluating the association of *CDH1* genotypes and haplotypes with Gleason grade in PCA patients.

RESULTS

A description of the clinical populations studied is provided in Table I. Significant differences in mean age between cases but not between controls are observed, with European American cases being about 5 years younger on average than African American and Jamaican cases ($P < 0.001$). Mean age differs significantly between cases and controls only in European Americans and the combined sample ($P \leq 0.05$). There also appears to be substantial differences in the mean PSA between the three samples. African American cases have statistically higher PSA values than European American cases; in addition, Jamaican PSA values tended to be significantly higher than African Americans ($P < 0.001$). The significant differences in PSA levels between each group suggest that we may potentially be comparing different PCA disease phenotypes with respect to stage, however, the percentage of cases with Gleason grade ≥7 ranged from 44% among African Americans to 65% among Jamaicans.

We identified 21 polymorphisms in *CDH1*: 8 in the promoter, 9 in introns, and 4 in exons. Table II shows the polymorphisms that were found in the promoter, coding, and non-coding regions of *CDH1* and the putative alterations in transcription factor binding sites or amino acids due to the sequence variant. Three of the promoter SNPs have been described before; −160C/A by Li et al. [13] and Nakamura et al. [14], while −1004A/T and −906C/A can be found in dbSNP and HapMap websites. All other promoter SNPs are novel findings. In addition, we found four new intronic polymorphisms in introns 1, 3, 9, and 15.

We selected four *CDH1* variants to genotype in all populations. The promoter SNP −1004A/T was chosen because the T allele eliminates a C/EBPα and a Hb binding sites and creates a MEB-1 binding site. Another promoter SNP, −160C/A, was selected based on previous literature reports that describe its association with several types of cancers, and because it showed a minor allele frequency of ∼20% in our screening sample. In addition, intronic polymorphisms IVS1 + 5C/G and IVS1 + 6T/C were typed because they had a moderate to high minor allele frequency (9 and 16%, respectively), and because they could potentially affect splicing sites due to their proximity to the exon/intron boundary. Only SNP IVS1 + 6T/C deviated markedly from Hardy–Weinberg equilibrium among Jamaican patients ($P = 0.002$). Polymorphisms −1004A/T and −160C/A were moderately out of Hardy–Weinberg equilibrium among African American cases and European American controls, respectively ($P < 0.05$).

Genotype and variant allele frequencies of the four selected polymorphisms in the study populations are shown in Table III. Significant differences between populations were detected for SNPs −1004A/T, IVS1 + 5C/G, and IVS1 + 6T/C. Within each ethnicity, however, there were no differences in prevalence of allele between patients and controls. On the other hand, SNP −160C/A did not show any difference in allele

TABLE II. Polymorphisms Identified in the *CDH1* Screening of African American Controls

Position[a]	Location	Polymorphism	Frequency[b]	Effect[c]	Ref./dbSNP rs#
−1004	Promoter	A > T	0.07	C/EBPα and Hb → MEB-1	rs13335980
−906	Promoter	C > A	0.15	No change	rs7194355
−782	Promoter	C > T	0.01	Ap-2α → NF-1	This study
−752	Promoter	C > T	0.01	C/EBPα → C/EBPβ	This study
−599	Promoter	C > T	0.15	No change	This study
−486	Promoter	A insertion	0.10	Deletes C/EBPα	This study
−479	Promoter	G > T	0.02	No change	This study
−160	Promoter	C > A	0.19	68% decreased transcriptional activity	[14,38] rs16260
IVS1 + 5 + 53	Intron 1	C > G	0.09	n/a	This study
IVS1 + 6 + 54	Intron 1	T > C	0.16	n/a	[39] rs3743674
+110	Intron 1	13 bp deletion	0.10	n/a	[40] rs3833051
+123	Intron 1	7 bp insertion	0.05	n/a	[40] rs3833051
IVS3 + 76	Intron 3	C > A	0.05	n/a	This study
IVS4 + 541	Intron 4	G > C	0.04	n/a	[41]
+933	Exon 7	C > G	0.07	Leu311Leu	[42]
IVS9 + 45	Intron 9	G > C	0.13	n/a	This study
+1849	Exon 12	G > A	0.02	Ala617Thr	[42]
+1896	Exon 12	C > T	0.05	His632His	[42]
IVS12−13	Intron 12	T > C	0.05	n/a	[43]
+2253	Exon 14	C > T	0.24	Asn751Asn	[42]
IVS15 + 22	Intron 15	C > T	0.07	n/a	This study

SNPs typed in the association study are shown in boldface. n/a, not available.
[a]Number of base pairs from ATG start codon.
[b]Frequency of polymorphism in 24 African American unaffected controls.
[c]Transcriptional binding site change or amino acid change.

frequency across populations but European Americans cases and controls were significantly different in allele and genotype frequencies ($P < 0.05$). Jamaican affected and unaffected subjects differed significantly with respect to marker IVS1 + 6T/C genotypic frequencies ($P = 0.02$). Despite European American patients having higher frequencies of the −160A allele compared to controls of the same ancestry, the logistic regression analysis yielded a non-significant result for the age-adjusted analysis of −160C/A genotype and PCA risk (Table IV). The fact that the unadjusted genotype, allele and linear trend tests revealed modestly significant associations with PCA in European Americans is likely due to the absence of controls who were homozygote for the polymorphism. For the Jamaican subjects, SNP IVS1 + 6T/C was significantly associated with PCA even after controlling for age (Table IV). However, this result should be interpreted with caution as genotypes among Jamaican patients were noticeably out of Hardy–Weinberg equilibrium. Finally, no association with PCA was observed for SNPs −1004A/T and IVS1 + 5C/G in any population (Table IV).

Linkage disequilibrium between pairs of markers was quite strong across the region in all populations, especially between SNPs −1004A/T, IVS1 + 5C/G, and IVS1 + 6T/C ($P < 0.05$). Polymorphism −160C/A was not significantly linked to −1004A/T in any population, while it showed significant linkage only to SNP IVS1 + 6T/C in European Americans, and to IVS1 + 5C/G in African Americans. SNPs IVS1 + 5C/G and IVS1 + 6T/C were strongly linked in all groups due to the very small distance between them (1 base pair) (Table V).

Haplotype analysis revealed a greater number of haplotypes in the African American (eight haplotypes, five of them common, i.e., >5%) and Jamaican populations (nine haplotypes, four of them common), than in the European American population (six haplotypes, three of them common). Three haplotypes (ACCT, ACCC, and AACT) comprised ∼98% of all chromosomes in European Americans, but only ∼84% in African Americans and ∼77% in Jamaicans.

Among all cases combined and controls combined, we observed a marginally significant increased disease risk for haplotype AACT (OR = 1.5, 95%CI: 1.0–2.2), $P = 0.05$, Table VI). Subset analyses revealed no significant differences in the distribution of haplotypes among African American ($P = 0.79$) or Jamaican

TABLE III. Genotype and Allele Frequencies (%) of *CDH1* Polymorphisms in Prostate Cancer Cases and Controls

	All subjects		African Americans		European Americans		Jamaicans	
SNP	Cases (N = 427)	Controls (N = 337)	Cases (N = 119)	Controls (N = 112)	Cases (N = 219)	Controls (N = 102)	Cases (N = 89)	Controls (N = 123)
−1004 A/T								
AA	89.6	84.8	88.1	84.7	98.6	100.0	69.7	72.4
AT	9.5	14.0	10.2	15.3	1.4	0.0	28.1	24.4
TT	0.9	1.2	1.7	0.0	0.0	0.0	2.2	3.2
T	5.7	8.2	6.8	7.7	0.7	0.0	16.3	15.5
−160 C/A								
CC	62.2	66.3	68.6	68.4	51.1	60.7	79.0	69.4
CA	33.4	31.9	29.4	29.6	42.9	39.3	17.3	27.5
AA	4.4	1.8	2.0	2.0	6.0	0.0	3.7	3.1
A	21.1	17.7	16.7	16.8	27.5	19.7	12.3	16.8
IVS1 + 5 C/G								
CC	84.6	77.7	79.5	73.6	97.0	99.0	62.9	64.2
CG	14.4	19.9	17.9	24.6	3.0	1.0	36.0	30.8
GG	1.0	2.4	2.6	1.8	0.0	0.0	1.1	5.0
G	8.2	12.4	11.5	14.1	1.5	0.5	19.1	20.4
IVS1 + 6 T/C								
TT	56.1	50.4	51.3	50.0	74.8	67.0	20.2	37.5
TC	36.8	40.4	38.4	40.0	22.8	32.0	66.3	47.5
CC	7.1	9.2	10.3	10.0	2.5	1.0	13.5	15.0
C	25.5	29.4	29.5	30.0	13.9	17.0	46.6	38.8

Genotype and allele frequencies of −1004, IVS1 + 5, and IVS1 + 6 differ significantly between population controls ($P < 0.001$).
IVS1 + 5 genotype frequencies differ significantly between cases and controls when all populations are combined ($P = 0.04$).
SNP −160 genotype and allele frequencies differ significantly between cases and controls in European Americans ($P = 0.04$).
IVS1 + 6 genotype frequencies differ significantly between cases and controls in Jamaicans ($P = 0.02$).

($P = 0.69$) cases and controls. Interestingly, among European Americans, PCA patients and unaffected subjects differed significantly with respect to haplotype frequencies ($P = 0.02$). European American PCA cases displayed a higher frequency of the single common haplotype that carried the −160A variant (AACT), and a lower frequency of the wild-type haplotype (ACCT) compared to controls (Table VI). Haplotype AACT was associated with ∼2-fold increased risk for PCA in an unadjusted analysis in European Americans ($P = 0.04$). However, after stratifying the sample by age, the effect of the −160A allele on PCA susceptibility was more apparent among individuals under the age of 65 (OR = 3.2, 95% CI: 1.0–10.7, $P = 0.05$) than in older subjects. In Jamaicans, on the other hand, the association of SNP IVS1 + 6T/C with disease was stronger in the ≥60 group (OR = 2.7, 95%CI: 1.3–5.7, $P = 0.01$).

When we stratified all PCA cases according to disease aggressiveness (low grade vs. high grade Gleason score) we did not observe any correlation between disease aggression and individual SNP genotypes or haplotypes in any population (data not shown). We do note however, the limitations of stratification of PCA by Gleason grade. These include the potential within and between-observer variation in grading, especially across international sites [26–30]. Nevertheless, European Americans with more aggressive disease (Gleason score ≥7) showed a significantly higher −160A allele frequency than controls (0.29 vs. 0.20, respectively, $P = 0.04$).

DISCUSSION

Since Li et al. [13] identified the functional *CDH1* −160C/A SNP (68% decreased transcription for the A allele), a number of studies have evaluated the effect of this polymorphism on the development and progression of different types of cancer. In a case-control study of gastric cancer among Taiwanese conducted by Wu et al. [44], individuals homozygous for the A variant were five times less likely to contract the disease than those homozygous for the C wild-type allele. Additionally, the authors did not observe any correlation of *CDH1* genotype and tumor stage or lymph node

TABLE IV. *CDH1* Genotypes and Risk of Prostate Cancer

SNP genotype	OR[a]	95% CI	*P*-value
−1004 A/T			
AA	1.00	Reference	
All subjects AT/TT	0.93	0.58–1.49	0.75
African Americans AT/TT	0.76	0.35–1.64	0.49
European Americans AT/TT[b]	—	—	—
Jamaicans AT/TT	1.00	0.54–1.88	0.99
−160 C/A			
CC	1.00	Reference	
All subjects CA/AA	0.92	0.62–1.35	0.66
African Americans CA/AA	1.07	0.58–1.98	0.82
European Americans CA/AA	1.23	0.57–2.64	0.60
Jamaicans CA/AA	0.57	0.28–1.15	0.12
IVS1 + 5 C/G			
CC	1.00	Reference	
All subjects CG/GG	0.86	0.57–1.31	0.49
African Americans CG/GG	0.75	0.40–1.40	0.37
European Americans CG/GG	0.56	0.06–5.45	0.62
Jamaicans CG/GG	1.01	0.56–1.81	0.99
IVS1 + 6 T/C			
TT	1.00	Reference	
All subjects TC/CC	1.22	0.85–1.76	0.28
African Americans TC/CC	0.98	0.58–1.66	0.94
European Americans TC/CC	0.78	0.35–1.77	0.55
Jamaicans TC/CC	2.27	1.19–4.34	0.01

[a]OR adjusted by age and ethnicity in all subjects and by age in each population.
[b]This SNP was monomorphic in European American controls.

metastasis. In contrast, Humar et al. [34] described the association of the mutant A allele with sporadic diffuse gastric cancer in Italy. No association was found between −160C/A genotype and risk of stomach cancer in three populations of European origin, namely Canadians, Germans, and Portuguese [31]. In urothelial cancer, however, Japanese individuals carrying the AA genotype had a 2.3-fold increased risk of being affected by the disease than CC individuals, although no correlation with tumor progression was detected [32]. An association of the A allele with risk for transitional cell carcinoma of the bladder as well as its correlation with malignancy progression was observed among Chinese [33]. Regarding PCA, conflicting results have been published. Elevated risk for Dutch A-carriers was initially described by Verhage et al. [16], with a stronger effect among sporadic cancer cases (∼5-fold) than among hereditary cancer patients (∼2-fold). An opposite finding was reported by Jonsson et al. [17] using pooled studies from Sweden. There they observed an association between hereditary PCA and the −160C/A genotype so that A-carriers were twice more likely to develop the disease than CC subjects. It should be noted that the association was not observed for sporadic PCA. The effect of the A allele was found to be even lower and non-significant, among sporadic cases from a Slovenian population (OR = 1.4, 95%CI = 0.9–2.4) [18]. In addition, no significant correlation of genotype and stage of disease was observed in our study. Similarly, risk of PCA, as well as tumor invasiveness and differentiation were not found to be associated with *CDH1* −160C/A polymorphism in Japanese by Tsukino et al. [20]. On the other hand, Kamoto et al. [19] reported a significant association between advanced prostate cancer and −160C/A genotype (CA + AA) compared to male controls in Japanese.

In this study, we have identified new polymorphisms in the promoter and introns of the *CDH1* gene amid others already described. In addition to SNP −160C/A we have genotyped three of these variants in a diverse population of cases and controls consisting of African Americans, European Americans, and Jamaicans. Genotype frequency differences between affected and unaffected individuals were apparent only for SNPs −160C/A and IVS1 + 6T/C in European Americans and Jamaicans, respectively. Particularly, the absence of individuals homozygous for the −160A

TABLE V. Pairwise Linkage Disequilibrium (D′) Between *CDH1* SNPs by Population

African Americans	−1004	−160	IVS1+5	IVS1+6
−1004A/T	—	0.308	**0.505**	*0.422*
−160C/A	0.580	—	0.160	0.212
IVS1 + 5C/G	**0.451**	**0.636**	—	**0.685**
IVS1 + 6T/C	**0.497**	0.286	**0.562**	—
European Americans	−1004	-160	IVS1+5	IVS1+6
−1004A/T	—	1.000	*0.154*	*0.613*
−160C/A	n/a	—	0.095	*0.475*
IVS1 + 5C/G	n/a	1.000	—	**0.807**
IVS1 + 6T/C	n/a	0.003	**1.000**	—
Jamaicans	−1004	-160	IVS1+5	IVS1+6
−1004A/T	—	0.221	**0.425**	**0.289**
−160C/A	0.152	—	0.325	0.307
IVS1 + 5C/G	**0.442**	0.273	—	**0.421**
IVS1 + 6T/C	**0.365**	0.220	**0.534**	—

Cases and controls are depicted above and below the diagonal, respectively. In bold, significant D′ ($P < 0.05$) based on the exact probability test. In italics, significant D′ ($P < 0.05$) based only on the Chi-square test.

allele among European American controls is noteworthy, as is the elevated frequency of this variant among subjects with aggressive disease and among younger patients (∼30%). Therefore, the impact of this promoter variant may be more relevant in the progression to aggressive disease rather than in the development of PCA per se. None of the remaining polymorphisms were significantly correlated with PCA except for marker IVS1 + 6T/C in Jamaicans. The association among Jamaicans should be considered with caution since IVS1 + 6T/C was noticeably out of Hardy–Weinberg equilibrium.

The frequency of the A allele in European American controls was lower than that reported for most European populations in other PCA studies [16–18] or non-PCA studies [31,34], but was comparable to Japanese population frequencies [20,32]. With the exception of SNP −160C/A, we found significant differences between ethnicities in genotype and allele frequencies of all *CDH1* markers.

Our findings are intriguing given previous findings for −160C/A and PCA risk. While two studies have shown low to moderate effects of the −160 A allele on hereditary prostate cancer (HPC), there has not been a consensus about its involvement in sporadic PCA [16,17]. Given that our study consisted of mainly sporadic PCA patients we could not evaluate the role of *CDH1* variants on HPC.

Downregulation of E-cadherin expression has been shown to occur during initiation and progression of PCA [35–37], however, it is not yet clear which role, if any, promoter polymorphisms play on the levels of protein production. Our results support an effect of the −160C/A promoter polymorphism in European Americans, although it is likely that such an effect is due to linkage with other variant or variants. If so, dissimilar levels of linkage disequilibrium between marker −160C/A and an unknown causal SNP in different populations could explain why an association is seen in certain groups and not in others. In fact, analysis of HapMap data for *CDH1* shows a different block structure for European Americans, Japanese, Chinese, and Yoruba, with the former exhibiting more extensive blocks and stronger linkage between blocks (data not shown). Two of the additional SNPs tested here (−1004A/T and IVS1 + 5C/G) are not linked to −160C/A in European Americans and therefore do not show any correlation with PCA. In spite of SNP IVS1 + 6 being tightly linked to −160C/A no association between the marker and disease was evident. However, when a haplotype-trait association test was performed, haplotype AACT appeared to confer increased risk for PCA. Similar findings were reported for gastric cancer by Humar et al. [34], who also typed the IVS1 + 6T/C polymorphism and an additional SNP in exon 13 (2076C/T). These researchers proposed that haplotype −160A, IVS1 + 6T, 2076T be considered as a marker for diffuse gastric cancer susceptibility within the Italian region sampled. Haplotype −160C, IVS1 + 6T, 2076T, was found to be protective in that study. So, according to Humar et al. [34] and our results, it seems that the presence of the A variant at position −160 together with the T allele at position +54 from the transcription start site, is associated with susceptibility to cancer in European populations. Conversely, when the wild-type −160C allele is present, the CT (−160, IVS1 + 6) haplotype decreases the risk of disease. Our results are insightful and

TABLE VI. Estimated *CDH1* Haplotype Frequencies in Prostate Cancer Cases and Controls and Disease Risk by Population

Population/haplotypes	Cases (%)	Controls (%)	OR[a]	95% CI	*P*-value
All subjects					
ACCT	54.5	55.3	0.98	0.73–1.30	0.86
ACCC	16.7	14.7	1.14	0.77–1.70	0.50
ACGC	2.2	4.4	0.46	0.20–1.07	0.07
AACT	19.9	14.4	1.46	0.99–2.15	0.05
TCGC	4.8	6.8	0.67	0.36–1.24	0.20
Rare haplotypes	1.9	4.4	—	—	—
$P = 0.13$[b]					
African Americans					
ACCT	55.0	56.0	0.94	0.56–1.57	0.80
ACCC	17.7	13.5	1.39	0.67–2.85	0.37
ACGC	3.8	5.9	0.66	0.20–2.14	0.48
AACT	14.9	13.6	1.15	0.55–2.41	0.71
TCGC	5.8	7.6	0.72	0.26–1.99	0.52
Rare haplotypes	2.8	3.4	—	—	—
$P = 0.79$[b]					
European Americans					
ACCT	59.2	66.8	0.73	0.45–1.19	0.21
ACCC	11.9	13.0	0.92	0.45–1.88	0.82
ACGC	0.4	0.5	0.46	0.03–7.48	0.58
AACT	26.9	16.2	1.84	1.01–3.36	0.04
TCGC	0.2	0.0	—	—	—
Rare haplotypes	1.4	3.5	—	—	—
$P = 0.02$[b]					
Jamaicans					
ACCT	42.6	45.8	0.89	0.51–1.54	0.68
ACCC	26.4	16.4	1.79	0.91–3.52	0.09
ACGC	3.4	6.2	0.50	0.13–1.95	0.31
AACT	10.2	13.0	0.75	0.32–1.79	0.52
TCGC	14.6	12.0	1.23	0.55–2.74	0.61
Rare haplotypes	2.8	6.6	—	—	—
$P = 0.69$[b]					

[a]OR calculated testing each haplotype against all other haplotypes combined.
[b]*P*-value for the omnibus test of differences in haplotype frequencies between cases and controls.

suggest that E-cadherin may contribute to PCA risk in a complex manner due to multiple variants, each of which may exhibit differing effects. However, in the future, other polymorphisms and haplotypes in the region should be evaluated in order to quantify predisposing polymorphisms and the at-risk haplotypic background.

ELECTRONIC DATABASE INFORMATION

Prediction of transcription factor binding sites in the promoter region of the *CDH1* gene was performed using a web-based program: http://www.generegulation.com

Online Mendelian Inheritance in Man (OMIM): http://www.ncbi.nih.gov/OMIM [for *CDH1* (MIM 192090), *CYP3A4* (MIM 124010), *CYP3A5* (MIM 605325), *SRD5A2* (MIM 607306), and *VDR* (MIM 601769)]

dbSNP for SNP information: http://www.ncbi.nlm.nih.gov/dbSNP

HapMap: http://www.hapmap.org/

ACKNOWLEDGMENTS

The authors would like to thank all the men who volunteered to participate in this genetic study.

REFERENCES

1. Siddiqui E, Mumtaz F, Gelister J. Understanding prostate cancer. J R Soc Health 2004;124(5):219–221.

2. Simard J, Dumont M, Labuda D, Sinnett D, Meloche C, El-Alfy M, Berger L, Lees E, Labrie F, Tavtigian S. Prostate cancer susceptibility genes: Lessons learned and challenges posed. Endocr Relat Cancer 2003;10(2):225–259.
3. Gsur A, Feik E, Madersbacher S. Genetic polymorphisms and prostate cancer risk. World J Urol 2004;21(6):414–423.
4. Pecina-Slaus N. Tumor suppressor gene E-cadherin and its role in normal and malignant cells. Cancer Cell Int 2003;3(1):17.
5. Hazan R, Qiao R, Keren R, Badano I, Suyama K. Cadherin switch in tumor progression. Ann NY Acad Sci 2004;1014:155–163.
6. Tsutsumida A, Hamada J, Tada M, Aoyama T, Furuuchi K, Kawai Y, Yamamoto Y, Sugihara T, Moriuchi T. Epigenetic silencing of E- and P-cadherin gene expression in human melanoma cell lines. Int J Oncol 2004;25(5):1415–1421.
7. Lin Y, Wu M, Li D, Wu X, Zheng R. Prognostic and clinicopathological features of E-cadherin, alpha-catenin, beta-catenin, gamma-catenin and cyclin D(1) expression in human esophageal squamous cell carcinoma. World J Gastroenterol 2004;10(22):3235–3239.
8. Salon C, Moro D, Lantuejoul S, Brichon Py P, Drabkin H, Brambilla C, Brambilla E. E-cadherin-beta-catenin adhesion complex in neuroendocrine tumors of the lung: A suggested role upon local invasion and metastasis. Hum Pathol 2004;35(9):1148–1155.
9. Oliveira C, Ferreira P, Nabais S, Campos L, Ferreira A, Cirnes L, Castro-Alves C, Veiga I, Fragoso M, Regateiro F, Moreira Dias L, Moreira H, Suriano G, Machado J, Lopes C, Castedo S, Carneiro F, Seruca R. E-cadherin (CHD1) and p53 rather than SMAD4 and Caspase-10 germline mutations contribute to genetic predisposition in Portuguese gastric cancer patients. Eur J Cancer 2004;40(12):1897–1903.
10. Suriano G, Oliveira M, Huntsman D, Mateus A, Ferreira P, Casares F, Oliveira C, Carneiro F, Machado J, Mareel M, Seruca R. E-cadherin germline missense mutations and cell phenotype: Evidence for the independence of cell invasion on the motile capabilities of the cells. Hum Mol Genet 2003;12(22):3007–3016.
11. Sarrio D, Moreno-Bueno G, Hardisson D, Sanchez-Estevez C, Guo M, Herman J, Gamallo C, Esteller M, Palacios J. Epigenetic and genetic alterations of *APC* and *CDH1* genes in lobular breast cancer: Relationships with abnormal E-cadherin and catenin expression and microsatellite instability. Int J Cancer 2003; 106(2):208–215.
12. Wheelock M, Johnson K. Cadherins as modulators of cellular phenotype. Annu Rev Cell Dev Biol 2003;19:207–235.
13. Li L, Zhao H, Nakajima K, Oh B, Alves Ribeiro Filho L, Carroll P, Dahiya R. Methylation of the E-cadherin gene promoter correlates with progression of prostate cancer. J Urol 2001; 166(2):705–709.
14. Nakamura A, Shimazaki T, Kaneko K, Shibata M, Matsumura T, Nagai M, Makino R, Mitamura K. Characterization of DNA polymorphisms in the E-cadherin gene (*CDH1*) promoter region. Mutat Res 2002;502(1–2):19–24.
15. Berx G, Becker K, Hofler H, van Roy F. Mutations in the human E-cadherin (*CDH1*) gene. Hum Mutat 1998;12(4):226–237.
16. Verhage B, van Houwelingen K, Ruijter T, Kiemeney L, Schalken J. Single-nucleotide polymorphism in the E-cadherin gene promoter modifies the risk of prostate cancer. Int J Cancer 2002;100(6):683–685.
17. Jonsson B, Adami H, Hagglund M, Bergh A, Goransson I, Stattin P, Wilklund F, Gronberg H. −160C/A polymorphism in the E-cadherin gene promoter and risk of hereditary, familial and sporadic prostate cancer. Int J Cancer 2004;109(3):348–352.
18. Hajdinjak T, Toplak N. E-cadherin polymorphism −160C/A and prostate cancer. Int J Cancer 2004;109(3):480–481.
19. Kamoto T, Isogawa Y, Shimizu Y, Minamiguchi S, Kinoshita H, Kakehi Y, Mitsumori K, Yamamoto S, Habuchi T, Kato T, Ogawa O. Association of a genetic polymorphism of the E-cadherin gene with prostate cancer in a Japanese population. Jpn J Clin Oncol 2005;35(3):158–161.
20. Tsukino H, Kuroda Y, Imai H, Nakao H, Qiu D, Komiya Y, Inatomi H, Hamasaki T, Kohshi K, Osada Y, Katoh T. Lack of evidence for the association of E-cadherin gene polymorphism with increased risk or progression of prostate cancer. Urol Int 2004;72(3):203–207.
21. Panguluri R, Long L, Chen W, Wang S, Coulibaly A, Ukoli F, Jackson A, Weinrich S, Ahaghotu C, Isaacs W, Kittles R. COX-2 gene promoter haplotypes and prostate cancer risk. Carcinogenesis 2004;25(6):961–966.
22. Grabe N. AliBaba2: Context specific identification of transcription factor binding sites. In Silico Biol 2002;2(1):S1–S15.
23. Hawley M, Kidd K. Haplo: A program using the EM algorithm to estimate the frequencies of multi-site haplotypes. J Hered 1995;86:409–411.
24. Long J, Williams R, Urbanek M. An E-M algorithm and testing strategy for multiple-locus haplotypes. Am J Hum Genet 1995; 56(3):799–810.
25. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 1995;12:921–927.
26. Bostwick D. Gleason grading of prostatic needle biopsies: Correlation with grade in 316 matched prostatectomies. Am J Surg Pathol 1994;18(8):796–803.
27. Steinberg D, Sauvageot J, Piantadosi S, Epstein J. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. Am J Surg Pathol 1997; 21(5):566–576.
28. Carlson G, Calvanese C, Kahane H, Epstein J. Accuracy of biopsy Gleason scores from a large uropathology laboratory: Use of a diagnostic protocol to minimize observer variability. Urology 1998;51(4):525–529.
29. Freeman V, Coard K, Wojcik E, Durazo-Arvizu R. Use of the Gleason system in international comparisons of prostatic adenocarcinomas in Blacks. Prostate 2004;58(2):169–173.
30. Coard K, Freeman V. Gleason grading of prostate cancer. Am J Clin Pathol 2004;122(3):373–376.
31. Pharoah P, Oliveira C, Machado J, Keller G, Vogelsang H, Laux H, Becker K, Hahn H, Paproski S, Brown L, Caldas C, Huntsman D. CDH1 c-160a promotor polymorphism is not associated with risk of stomach cancer. Int J Cancer 2002;101(2):196–197.
32. Tsukino H, Kuroda Y, Nakao H, Imai H, Inatomi H, Kohshi K, Osada Y, Katoh T. E-cadherin gene polymorphism and risk of urothelial cancer. Cancer Lett 2003;195(1):53–58.
33. Zhang X, Ma X, Zhu Q, Li L, Chen Z, Ye Z. Association between a C/A single nucleotide polymorphism of the E-cadherin gene promoter and transitional cell carcinoma of the bladder. J Urol 2003;170(4 Pt 1):1379–1382.
34. Humar B, Graziano F, Cascinu S, Catalano V, Ruzzo A, Magnani M, Toro T, Burchill T, Futschik M, Merriman T, Guilford P. Association of CDH1 haplotypes with susceptibility to sporadic diffuse gastric cancer. Oncogene 2002;21(53):8192–8195.
35. Rubin M, Mucci N, Figurski J, Fecko A, Pienta K, Day M. E-cadherin expression in prostate cancer: A broad survey using high-density tissue microarray technology. Hum Pathol 2001;32(7):690–697.

36. Kallakury B, Sheehan C, Winn-Deen E, Oliver J, Fisher H, Kaufmann RJ, Ross J. Decreased expression of catenins (alpha and beta), p120CTN, and E-cadherin cell adhesion proteins and E-cadherin gene promoter methylation in prostatic adenocarcinomas. Cancer 2001;92(11):2786–2795.
37. Wehbi N, Dugger A, Bonner R, Pitha J, Hurst R, Hemstreet Gr. Pan-cadherin as a high level phenotypic biomarker for prostate cancer. J Urol 2002;167(5):2215–2221.
38. Li L, Chui R, Sasaki M, Nakajima K, Perinchery G, Au H, Nojima D, Carroll P, Dahiya R. A single nucleotide polymorphism in the e-cadherin gene promoter alters transcriptional activities. Cancer Res 2000;60:873–876.
39. Avizienyte E, Launonen V, Salovaara R, Kiviluoto T, Aaltonen L. E-cadherin is not frequently mutated in hereditary gastric cancer. J Med Genet 2001;38(1):49–52.
40. Haga H, Yamada R, Ohnishi Y, Nakamura Y, Tanaka T. Gene-based SNP discovery as part of the Japanese Millennium Genome Project: Identification of 190,562 genetic variations in the human genome. Single nucleotide polymorphism. J Hum Genet 2002;47(11):605–610.
41. Berx G, Cleton-Jansen A, Nollet F, de Leeuw W, van de Vijver M, Cornelisse C, van Roy F. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. EMBO J 1995;14(24):6107–6115.
42. Risinger J, Berchuck A, Kohler M, Boyd J. Mutations of the E-cadherin gene in human gynecologic cancers. Nat Genet 1994;7(1):98–102.
43. Berx G, Cleton-Jansen A, Strumane K, de Leeuw W, Nollet F, van Roy F, Cornelisse C. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene 1996;13(9):1919–1925.
44. Wu MS, Huang SP, Chang YT, Lin MT, Shun CT, Chang MC, Wang HP, Chen CJ, Lin JT. Association of the −160C→A promoter polymorphism of E-cadherin gene with gastric carcinoma risk. Cancer 2002;94(5):1443–1448.