NIH Public Access
Author Manuscript
Prostate. Author manuscript; available in PMC 2010 March 1.

Published in final edited form as:
Prostate. 2009 March 1; 69(4): 363–372. doi:10.1002/pros.20887.

Clinical Utility of Five Genetic Variants for Predicting Prostate Cancer Risk and Mortality

Claudia A. Salinas[1,2], **Joseph S. Koopmeiners**[1,3], **Erika M. Kwon**[4], **Liesel FitzGerald**[1], **Daniel W. Lin**[5], **Elaine A. Ostrander**[4], **Ziding Feng**[1], and **Janet L. Stanford**[1,2]

[1] Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, Seattle, WA

[2] Department of Epidemiology, School of Public Health and Community Medicine, University of Washington, Seattle, WA

[3] Department of Biostatistics, School of Public Health and Community Medicine, University of Washington, Seattle, WA

[4] Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD

[5] Department of Urology, University of Washington Medical Center, Seattle, WA

Abstract

Background—A recent report suggests that the combination of five single-nucleotide polymorphisms (SNPs) at 8q24, 17q12, 17q24.3 and a family history of the disease may predict risk of prostate cancer. The present study tests the performance of these factors in prediction models for prostate cancer risk and prostate cancer-specific mortality.

Methods—SNPs were genotyped in population-based samples from Caucasians in King County, Washington. Incident cases (n=1308), aged 35–74, were compared to age-matched controls (n=1266) using logistic regression to estimate odds ratios (OR) associated with genotypes and family history. Cox proportional hazards models estimated hazard ratios for prostate cancer-specific mortality according to genotypes.

Results—The combination of SNP genotypes and family history was significantly associated with prostate cancer risk ($p_{trend}=1.5 \times 10^{-20}$). Men with ≥ five risk factors had an OR of 4.9 (95% CI 1.6 to 18.5) compared to men with none. However, this combination of factors did not improve the ROC curve after accounting for known risk predictors (i.e., age, serum PSA, family history). Neither the individual nor combined risk factors was associated with prostate cancer-specific mortality.

Conclusion—Genotypes for five SNPs plus family history are associated with a significant elevation in risk for prostate cancer and may explain up to 45% of prostate cancer in our population. However, they do not improve prediction models for assessing who is at risk of getting or dying from the disease, once known risk or prognostic factors are taken into account. Thus, this SNP panel may have limited clinical utility.

Keywords

prostate cancer; 8q24; 17q12; 17q24; screen; genetic

Requests for reprints: Janet L. Stanford, Program in Epidemiology, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, M4-B874, PO Box 19024, Seattle, WA 98109-1024., Phone: 206-667-2715; Fax: 206-667-4787; jstanfor@fhcrc.org.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Introduction

Prostate cancer is a common disease in the U.S., with an estimated probability of being diagnosed of 2.5%, 7% and 13% for men ages 40–59 years, 60–69 years and 70 years and older, respectively (1). Established risk factors for prostate cancer include increasing age, African ancestry, and a first-degree family history of prostate cancer. Family history can help identify men with a genetic predisposition to prostate cancer, but recent technological advances have allowed investigators to interrogate thousands of single nucleotide polymorphisms (SNPs) across the genome to search for specific genetic markers associated with risk of developing this complex disease.

Several recent linkage and genome-wide association scans have identified polymorphisms that influence risk of prostate cancer (2–7) and many of the recent findings from large association studies have been replicated in different populations. Even though the risk estimates for individual SNPs have been modest (5,8–13), this has led to interest in the possibility of developing a genetic test to predict risk. However, for a risk factor (e.g., biomarker) to function adequately as an effective screening test, it must be strongly associated with disease (14,15). For example, Pepe and colleagues demonstrated that for a binary marker to attain a test sensitivity of 80% (true positives) and a specificity of 90% (10% false positives), an OR of about at least 36 is required for the marker-disease association (15). None of the individual SNP associations reported to date has approached this magnitude of risk.

In an effort to overcome this limitation, a recent study from Sweden (16) examined a combination of five independent SNPs that map to 8q24, 17q12, and 17q24.3 (rs16901979, rs6983267, rs1447295, rs4430796, and rs1859962) plus family history in relation to risk of prostate cancer. Individually, the risk estimates for each SNP ranged from 1.2 to 1.5 in a multivariate model with all five SNPs. However, men who carried four or more of the high-risk genotypes (5.4% of cases, 2.2% of controls) had a risk estimate of 4.47 (95% CI 2.93 to 6.80), which increased to 9.46 (95% CI 3.62 to 24.72) in men with all five high-risk genotypes or four high-risk genotypes plus a positive family history (1.4% of cases, 0.3% of controls) relative to men with none of the high-risk genotypes and no family history. Initial confirmation of these results has been reported (17), with an 11-fold increase in the relative risk of prostate cancer for men with at least five of the six risk indicators (i.e., five SNP genotypes and family history). In the Swedish study (16), there was no evidence that any of the SNP genotypes was associated with features of more clinically aggressive prostate cancer such as tumor stage, Gleason score, PSA level at diagnosis; there was also no relationship with age at diagnosis or family history status. Nonetheless, based on the above results, cancer risk tests have been developed as tools to assess an individual's genetic risk of developing prostate cancer (e.g., the Focus5™ Prostate Cancer Risk Test www.proactivegenomics.com and the ProCa™ test, www.decodediagnostics.com).

The present study was designed to evaluate whether the findings of Zheng et al. (16) could be replicated in a population-based sample of American Caucasian men and to evaluate how the combination of SNP genotypes and family history function in prediction models for prostate cancer risk and for prostate cancer-specific mortality.

Subjects and Methods

Study Population

The study population consists of participants from two population-based case-control studies in residents of King County, Washington (Study I and Study II), which have been described previously (18,19). Incident Caucasian and African American cases with histologically

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

confirmed prostate cancer were ascertained from the Seattle-Puget Sound SEER cancer registry. In Study I, cases were diagnosed between January 1, 1993, and December 31, 1996 and were 40–64 years of age at diagnosis. In Study II, cases were diagnosed between January 1, 2002, and December 31, 2005 and were 35–74 years of age at diagnosis. Of the 2244 eligible prostate cancer patients identified, 1754 (78.2%) were interviewed. The main reasons for non-response were patient refusal (13.9%), physician refusal to allow patient contact (2.1%), patients were too ill to participate (0.9%), or died before interview (1.4%). DNA for genotyping was prepared from blood samples collected from 1457 (83%) interviewed cases using standard protocols (20).

Controls were male residents of King County, Washington without a self-reported history of prostate cancer. They were identified using one-step random digit telephone dialing (21) and frequency matched to cases by five-year age groups. Controls were recruited evenly throughout both ascertainment periods for case patients. Complete household census information was obtained for 94% and 81% of the residential telephone numbers contacted for Study I and Study II, respectively. Of the 2448 men identified who met the eligibility criteria, 1645 (67.2%) completed a study interview. The main reasons for non-participation included refusal (29.1%) or too ill to participate (1.4%). Blood samples were collected and DNA prepared from 1351 (82%) interviewed controls.

Subjects in both studies completed in-person interviews conducted by trained male interviewers using a standardized questionnaire. Questions pertained to the time prior to reference date, i.e., the date of prostate cancer diagnosis for cases and for controls, a pre-assigned random date that approximated the distribution of cases' diagnosis dates. The questionnaire collected information on social and demographic factors, family structure, cancer history and medical history. Clinical information on cases, including Gleason score, stage of disease, serum PSA level at diagnosis, and primary treatment was obtained from the cancer registry. For this analysis, only Caucasian cases (n=1308) and controls (n=1266) were included.

Cases from both studies are under long-term surveillance with ascertainment of vital status and underlying cause of death (prostate cancer or other causes). The patient file is linked to the SEER cancer registry and for each deceased case patient a death certificate is requested from the state where the patient died, to confirm the underlying cause of death. The date of last follow-up for vital status for this analysis was November 1, 2007, with an average length of follow-up of 7.6 years (11.9 years for Study I and 4.0 years for Study II). All study procedures were approved by the Fred Hutchinson Cancer Research Center and NHGRI institutional review boards and written informed consent was obtained from all study subjects before participation.

Genotyping

SNPs were genotyped using the Applied Biosystems (ABI) SNPlex™ Genotyping System. Identification of the specific SNP allele was carried out with the ABI 3730xl DNA Analyzer with GeneMapper® software used for allele assignment (www.appliedbiosystems.com). Quality control included genotyping of 140 blind duplicate samples distributed across all genotyping batches, with 100% agreement between duplicates for the five SNPs. Each batch of DNA aliquots genotyped incorporated similar numbers of case patient and control samples and laboratory personnel were blinded to the case-control status of samples.

Statistical Analyses

Departure from Hardy-Weinberg equilibrium for the five SNPs in Caucasian controls was assessed using Fisher's exact test. Pairwise linkage disequilibrium (LD) between SNPs was

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

also estimated based on r^2 statistics calculated in controls. The association between prostate cancer risk and either individual or cumulative SNP genotypes was estimated by odds ratios (OR) and 95% confidence intervals (95% CI) from unconditional logistic regression models (22). For each SNP genotype, models adjusted for age were used to test dominant, recessive, and additive (zero, one, or two copies of the associated allele) genetic models. The best-fitting genetic model for each SNP was selected based on the model with the greatest likelihood. The cumulative effect of SNP genotypes across the five SNPs on prostate cancer risk was tested by counting the number of high-risk genotypes in each subject. A similar analysis counted a positive family history of prostate cancer in first-degree relatives as one of the cumulative risk factors. Odds ratios for both analyses compared men with any of the prostate cancer associated risk factors to men with none. Confounding was evaluated by considering whether inclusion of other covariates changed the risk estimates by ≥ 10%. P-values were derived from likelihood ratio-based test statistics obtained by comparison of nested models. Statistically significant associations were those with a two-sided p-value <0.05. Linear patterns in ORs were evaluated by a two-sided Cochran-Armitage test for trend. Goodness-of-fit was evaluated with the Hosmer-Lemeshow test. The existence of gene-gene and gene-environment interaction was evaluated using the likelihood ratio test comparing the full model with main effects and an interaction term for the covariates of interest to the reduced model without the interaction term. Statistical calculations were performed using SAS (v9.1.3) and R (v2.6.0).

Population attributable risk percent (PAR%), the proportion of disease in the population that can be attributed to one or more exposures, was calculated for each SNP based on the OR obtained from multivariate models adjusted for the other SNPs, first-degree family history of prostate cancer and five-year age groups (23). The joint PAR% for all five SNPs and for the SNPs plus family history was calculated from the individual PAR%s for comparison with results from Zheng et al. (16) Corrected PAR%s take into account the inflation resulting from use of the ORs to estimate the relative risk for a disease that is not rare in the population. Corrected individual PAR% were calculated by solving a quadratic equation in which the absolute risk is a function of the observed OR(s), exposure prevalence in population controls, and background disease prevalence (estimated from the SEER 9 limited-duration prevalence database) (24) and corrected joint PAR%s were then calculated from the individual corrected PAR%s.

Cox proportional hazards models were used to estimate hazard ratios (HR) and 95% confidence intervals for prostate cancer-specific mortality (25). Cox models were adjusted for age, Gleason score (2–4, 5–6, 7=3+4, 7=4+3, 8–10), serum PSA at diagnosis, stage (local, regional, distant), and primary treatment (radical prostatectomy, radiation, androgen deprivation therapy, other treatments and active surveillance). The proportional hazards assumption was assessed by testing the significance of time-dependent covariates created from the interaction of the predictor and survival time, with any statistically significant results indicating that the predictors were not proportional and the models not valid. The interaction terms were used only to test that the hazard function was constant over survival time and were not included in the models presented.

For predicting prostate cancer risk, logistic regression models were fit using the clinical characteristics (age, serum PSA, and family history) with and without the five SNPs. This analysis was restricted to cases (n=475) and controls (n=364) from Study I with complete information for age, serum PSA (at diagnosis for cases and at interview for controls), family history of prostate cancer, and genotypes for the five SNPs. Empirical receiver operating characteristic (ROC) curves for ten-year prostate cancer-specific mortality were drawn using risk scores from the Cox proportional hazards regression for the two models containing clinical characteristics (age, stage, Gleason score, PSA at diagnosis, family history and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

primary treatment) with and without the five SNPs. The resulting ROC curves compare the ability of models with and without the five SNPs to accurately classify cases vs. controls for the risk model and among cases for ten-year prostate cancer-specific mortality. Improvement in diagnostic accuracy due to the addition of the five SNPs was summarized by the difference in the area under the curve (AUC). Confidence intervals for the difference in AUC were calculated using the nonparametric bootstrap.

Results

The mean age of cases was similar to that of controls, 59.9 and 59.6 years, respectively. In comparison with controls, a higher proportion of cases had a first-degree family history of prostate cancer and reported having a PSA test within the five-year period before reference date (Table I). The majority of cases had serum PSA values of 4.0–9.9 ng/mL at diagnosis, localized stage disease and Gleason scores of 5 or 6; most were treated with radical prostatectomy.

The SNPs evaluated were the five previously reported by Zheng et al. (16), except for rs16901979. Instead, we genotyped rs6983561, which is perfectly correlated with SNP rs16901979 (r^2=1, based on HapMap CEPH individuals) and captures identical information. All SNP frequencies observed in controls were consistent with Hardy-Weinberg equilibrium.

Table II presents risk estimates for each SNP adjusted for the other SNPs and for first-degree family history of prostate cancer. As shown, a first-degree family history of prostate cancer confers a greater risk than any of the individual SNP genotypes (OR=2.32, 95% CI 1.85 to 2.92) even after controlling for all five SNPs. Risk estimates are presented separately for cases and controls from Study I with complete information for age, serum PSA, family history and the five SNPs. Results from this latter analysis were used to generate the ROC curves in Figure 1. The genetic model presented for each SNP (dominant, recessive or additive) follows the identical scheme as Zheng et al. (16) to allow for direct comparisons, although slightly different results were obtained in the present study. In the Swedish study, the best-fitting genetic model was selected from the statistical model with the highest likelihood. Results were identical in the present study, with the exception of SNP rs1859962, for which the dominant genetic model, as opposed to the recessive genetic model of Zheng et al. (16), provided the best fit ($p_{dominant}$=2.70= × 10^{-4} vs. $p_{recessive}$=1.9 × 10^{-2}). Stratifying the SNP analyses by a first-degree family history of prostate cancer revealed no evidence for effect modification. There was also no evidence of gene-gene interaction between the SNPs.

The uncorrected PAR% estimates (Table III) indicate that individual SNPs may account for 5% to 20% of all prostate cancers in the King County population, while the joint PAR% for all five SNPs plus a family history is 47%. After correcting these estimates for inflation, the joint PAR% drops to 45.4% assuming a 3% background prevalence of prostate cancer in the population and to 45.0% for a 5% background prevalence (24).

The five SNPs have a cumulative effect when modeled together, with men who carry at least one of the at-risk alleles having a significantly increased risk of prostate cancer (OR_{any}=1.89, 95% CI 1.42 to 2.52). Table IV shows that there is a linear increase in risk of disease with increasing number of at-risk alleles compared to men who have none of the at-risk alleles (p_{trend} = 1.2 × 10^{-12}). When family history of prostate cancer is also considered, the cumulative effect on risk is magnified (p_{trend}=1.5 × 10^{-20}). Men who carry two or more risk factors have over a two-fold increase in risk compared to men with none ($OR_{two\ risk\ factors}$=2.25, 95% CI 1.63 to 3.13), and the risk estimates increase with each

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

additional factor up to an OR = 4.92 (95% CI 1.6 to 18.5) for men who have a family history of prostate cancer and carry at least four of the at-risk alleles or for men who carry all five at-risk alleles.

The ROC curve in Figure 1 provides information about the ability of this group of SNPs to identify prostate cancer cases relative to controls. The figure compares a model with age at reference date, serum PSA level (at diagnosis for cases, at interview for controls) and first-degree family history of prostate cancer to a model with the five SNPs added. This figure includes a subset of Study I cases (n=475) and a random sample of controls (n=364) with complete data for age, serum PSA level, family history of prostate cancer and genotypes for the five SNPs. No significant differences were detected between demographic or clinical characteristics of cases who were and were not included in this analysis. No significant differences existed for controls, either, except that those with serum PSA values had a mean age of 56.0 years compared to those without serum PSA information, whose mean age was 57.0 years. The area under the curve (AUC) is 0.63 and 0.66, respectively (the difference between the curves is 0.03, 95% CI −0.12 to +0.06). Based on this analysis, the SNP genotypes do not substantially improve the ability to accurately predict risk of prostate cancer. In addition, none of the five SNPs was significantly associated with serum PSA levels in controls (data not shown).

Table V presents the association of SNP genotypes with ten-year prostate cancer-specific mortality. None of the SNPs was associated with dying from metastatic prostate cancer, neither in the multivariate model presented, nor in models with single SNPs (data not shown). Further, there was no cumulative effect of the five SNPs on mortality (data not shown). The addition of the five SNPs to an ROC curve (Figure 2) constructed with standard clinical parameters used for predicting prognosis (age at diagnosis, serum PSA at diagnosis, Gleason score, and tumor stage) did not appreciably increase the AUC for ten-year prostate cancer-specific mortality. The difference between the model with and without the SNPs is −0.005 (95% CI −0.01 to +0.02), indicating no improvement in prediction ability based on knowledge of the SNP genotypes.

Discussion

In a recent report Zheng et al. (16) analyzed five SNPs previously associated with prostate cancer risk and showed that the cumulative effect of carrying multiple at-risk genotypes plus a family history of prostate cancer increased risk up to an OR of 9.46 (95% CI 3.6 to 24.7). They also reported that this combination of risk factors (genotypes at five SNPs and family history of prostate cancer) could explain 46% of all prostate cancer in the Swedish population. To determine how these same factors relate to risk in a U.S. Caucasian population, we genotyped samples from two population-based case-control studies. As in the Swedish study, the five SNPs were significantly associated with prostate cancer risk in a multivariate model adjusted for the other SNPs and for first-degree family history of prostate cancer. The magnitude of risk estimates for individual SNPs and the combination of SNPs with family history in the multivariate models were also similar to the estimates reported by Zheng et al. (16) However, the cumulative effect of at-risk genotypes in our population was lower than that observed in the Swedish population, with U.S. Caucasian men who carry four or five of the at-risk genotypes having an OR for prostate cancer of 3.4 (95% CI 1.9 to 6.1) as compared to the OR of 4.5 (95% 2.9 to 6.8) reported in the Swedish population. Our results also differed when first-degree family history of prostate cancer was considered. In our population, men with five or six of the risk factors (i.e., 5 SNPs and family history) have a 5-fold increase in risk compared to men with none (95% CI 1.58 to 18.53), compared to the 9.5-fold increase in risk observed among the Swedish population (16). Interestingly, in

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

both studies a family history of prostate cancer conferred an elevated risk that was of greater magnitude than any of the individual SNP genotypes.

To further evaluate the potential clinical utility of these SNP genotypes for prediction of an individual's risk for prostate cancer, we generated ROC curves in a subset of men from Study I with complete information for age, serum PSA, and first-degree family history of prostate cancer (Figure 1). The results demonstrate that the addition of these SNP genotypes contributes minimal information toward the accurate classification of prostate cancer case patient status once age, PSA level and family history of the disease are taken into account.

The proportion of all prostate cancer cases in the population that can be explained by the cumulative effect of genotypes at these five SNPs and family history (i.e., the joint PAR%) was similar between the Swedish population (46.3%) and our U.S. Caucasian population (46.7%). These PAR% estimates are based on the population prevalence of exposure and the relative risk, which is estimated by the ORs in both case-control studies. For rare diseases, the OR provides a good estimate of the relative risk, but prostate cancer is not rare in the U.S. (1,24) As the prevalence of disease increases in the underlying population, the OR will be further from the true relative risk (26,27). Thus, the resulting individual SNP PAR% estimates will be inflated and this will lead to a compounding of the inflation effect when they are combined to generate a cumulative PAR%. SNPs chosen from genome-wide association studies for further analysis also tend to have inflated ORs due to their selection as top candidates, leading to another source of bias in estimating the PAR%s. However, in our data we did not observe such selection bias in the five SNPs, consistent with the literature that this bias is negligible when the OR is at least moderate, e.g., OR >1.3 (28). Based on SEER data, the estimated prevalence of prostate cancer in U.S. men of the same ages as those in our study population sample was 3.14% (24). Thus, after adjustment for inflation, our summary PAR% estimate was reduced from 46.7% to 45.4%, assuming a population prevalence of prostate cancer of 3% (Table III). This suggests that the combination of genotypes in these five SNPs and family history may explain up to 45% of prostate cancer cases in our population.

One component of evaluating the potential clinical utility of a screening test is its ability to improve disease outcomes. Zheng et al. (16) previously showed that the genotypes in five SNPs do not correlate with stage of disease at diagnosis, Gleason score or PSA level at diagnosis, which suggests that this SNP panel will not uniquely identify those men at higher risk for more clinically aggressive phenotypes. To further evaluate this issue, we analyzed the relationship between these SNP genotypes plus family history and prostate cancer-specific mortality. In our dataset, none of the SNP genotypes was associated with prostate cancer-specific mortality. The ROC curves demonstrate no benefit of adding the SNP genotypes to existing clinical prognostic factors (age, Gleason score, stage, serum PSA level and primary treatment) for identifying men who will die from prostate cancer (Figure 2). However, an important limitation of this analysis is the small number of deaths (n=45).

On the basis of the association between these five SNPs and risk of prostate cancer, Zheng et al. (16) are developing a genetic test, the Focus5™ Prostate Cancer Risk Test. In addition, the ProCa™ genetic test developed by deCODE genetics includes these same five SNPs in its panel of eight SNPs. Although the SNP genotypes are associated with an elevated relative risk of prostate cancer, the magnitude of the risk estimates, even for the combination of multiple at-risk alleles, suggests that this panel of gene variants may be of limited efficacy as a genetic screening test. As mentioned previously, the strength of the association between a risk factor and disease must be extremely strong (14,15) for the factor to perform well as a screening test. Zheng and colleagues acknowledge that the genetic test they are developing may only be useful for identifying men at high risk (17), who represent less than 2% of the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

controls in the current study population (see Table IV). It is anticipated that screening and earlier detection in men at highest risk of prostate cancer will lead to better outcomes, but randomized trials will be required to determine the efficacy of this genetic test panel in predicting an individual's risk of developing prostate cancer or disease outcome. The SNP genotypes are associated equally with less aggressive and with more aggressive clinical features of prostate cancer, so a screening test with these SNPs is unlikely to help in the search for biomarkers of more aggressive prostate cancer phenotypes. Also, there was no association detected between the SNPs and serum PSA levels at diagnosis, neither by Zheng et al. (16) nor in the present study. Thus, some men classified as at-risk based on the SNP panel will not have an elevated PSA level that would prompt a recommendation for prostate biopsy. This leads to important clinical questions, such as whether a biopsy will be considered for men who "test positive" based on their combined at-risk SNP profile despite not having any other indications for biopsy. This may be premature given the lack of evidence that the SNP panel can identify men at increased risk of developing more aggressive forms of prostate cancer.

Over the past few years, the number of genetic variants associated with risk of prostate cancer has increased substantially and there has been rapid replication of initial findings in diverse populations, such as for the SNPs evaluated in this study. However, although results from the SNP genotypes evaluated in this study and by Zheng et al. (16) contribute to our understanding of genetic susceptibility to prostate cancer, these SNP genotypes alone may be of limited clinical value for predicting risk of developing prostate cancer or of developing more aggressive prostate cancer. Additional translational studies may ultimately reveal how these common genetic variants may be used clinically, such as for risk stratification and in communicating risk-based information to individuals interested in early detection and prostate cancer prevention. However, finding and validating clinically useful screening tests for more aggressive forms of prostate cancer, including genetic-based tests of underlying predisposition remains an elusive goal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the dedicated study managers and staff and all the men who participated in these studies without whose help this work would not be possible.

Funding: National Cancer Institute (RO1 CA56678, R01 CA092579, RO1 CA082664 to J.L.S., P50 CA097186 Pacific Northwest Cancer SPORE Program); Department of Defense (Training Grant PC061445 to C.A.S.), with additional support from the Fred Hutchinson Cancer Research Center and the Intramural Program of the National Human Genome Research Institute.

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer statistics. CA Cancer J Clin. 2008; 58(2):71–96. [PubMed: 18287387]
2. Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Cazier JB, Sainz J, Jakobsdottir M, Kostic J, Magnusdottir DN, Ghosh S, Agnarsson K, Birgisdottir B, Le Roux L, Olafsdottir A, Blondal T, Andresdottir M, Gretarsdottir OS, Bergthorsson JT, Gudbjartsson D, Gylfason A, Thorleifsson G, Manolescu A, Kristjansson K, Geirsson G, Isaksson H, Douglas J, Johansson JE, Balter K, Wiklund F, Montie JE, Yu X, Suarez BK, Ober C, Cooney KA, Gronberg H, Catalona WJ, Einarsson GV, Barkardottir RB, Gulcher JR, Kong A, Thorsteinsdottir U, Stefansson K. A common variant associated with prostate cancer in European and African populations. Nat Genet. 2006; 38(6):652–658. [PubMed: 16682969]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

3. 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. [PubMed: 16945910]

4. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, Rafnar T, Bergthorsson JT, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Xu J, Blondal T, Kostic J, Sun J, Ghosh S, Stacey SN, Mouy M, Saemundsdottir J, Backman VM, Kristjansson K, Tres A, Partin AW, Albers-Akkers MT, Godino-Ivan Marcos J, Walsh PC, Swinkels DW, Navarrete S, Isaacs SD, Aben KK, Graif T, Cashy J, Ruiz-Echarri M, Wiley KE, Suarez BK, Witjes JA, Frigge M, Ober C, Jonsson E, Einarsson GV, Mayordomo JI, Kiemeney LA, Isaacs WB, Catalona WJ, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet. 2007; 39(5):631–637. [PubMed: 17401366]

5. Wang L, McDonnell SK, Slusser JP, Hebbring SJ, Cunningham JM, Jacobsen SJ, Cerhan JR, Blute ML, Schaid DJ, Thibodeau SN. Two common chromosome 8q24 variants are associated with increased risk for prostate cancer. Cancer Res. 2007; 67(7):2944–2950. [PubMed: 17409399]

6. Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N, Wang Z, Welch R, Staats BJ, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Gelmann EP, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover R, Hunter DJ, Chanock SJ, Thomas G. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet. 2007; 39(5):645–649. [PubMed: 17401363]

7. Eeles RA, Kote-Jarai Z, Giles GG, Olama AA, Guy M, Jugurnauth SK, Mulholland S, Leongamornlert DA, Edwards SM, Morrison J, Field HI, Southey MC, Severi G, Donovan JL, Hamdy FC, Dearnaley DP, Muir KR, Smith C, Bagnato M, Ardern-Jones AT, Hall AL, O'Brien LT, Gehr-Swain BN, Wilkinson RA, Cox A, Lewis S, Brown PM, Jhavar SG, Tymrakiewicz M, Lophatananon A, Bryant SL, Horwich A, Huddart RA, Khoo VS, Parker CC, Woodhouse CJ, Thompson A, Christmas T, Ogden C, Fisher C, Jamieson C, Cooper CS, English DR, Hopper JL, Neal DE, Easton DF. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet. 2008; 40(3):316–321. [PubMed: 18264097]

8. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeney LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet. 2007; 39(8):977–983. [PubMed: 17603485]

9. Haiman CA, Patterson N, Freedman ML, Myers SR, Pike MC, Waliszewska A, Neubauer J, Tandon A, Schirmer C, McDonald GJ, Greenway SC, Stram DO, Le Marchand L, Kolonel LN, Frasco M, Wong D, Pooler LC, Ardlie K, Oakley-Girvan I, Whittemore AS, Cooney KA, John EM, Ingles SA, Altshuler D, Henderson BE, Reich D. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet. 2007; 39(5):638–644. [PubMed: 17401364]

10. Robbins C, Torres JB, Hooker S, Bonilla C, Hernandez W, Candreva A, Ahaghotu C, Kittles R, Carpten J. Confirmation study of prostate cancer risk variants at 8q24 in African Americans identifies a novel risk locus. Genome Res. 2007

11. Schumacher FR, Feigelson HS, Cox DG, Haiman CA, Albanes D, Buring J, Calle EE, Chanock SJ, Colditz GA, Diver WR, Dunning AM, Freedman ML, Gaziano JM, Giovannucci E, Hankinson SE, Hayes RB, Henderson BE, Hoover RN, Kaaks R, Key T, Kolonel LN, Kraft P, Le Marchand L, Ma J, Pike MC, Riboli E, Stampfer MJ, Stram DO, Thomas G, Thun MJ, Travis R, Virtamo J, Andriole G, Gelmann E, Willett WC, Hunter DJ. A common 8q24 variant in prostate and breast

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

cancer from a large nested case-control study. Cancer Res. 2007; 67(7):2951–2956. [PubMed: 17409400]

12. Salinas CA, Kwon E, Carlson CS, Koopmeiners JS, Feng Z, Karyadi DM, Ostrander EA, Stanford JL. Multiple independent genetic variants in the 8q24 region are associated with prostate cancer risk. Cancer Epidemiol Biomarkers Prev. 2008; 17(5):1203–1213. [PubMed: 18483343]
13. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, Yu K, Chatterjee N, Welch R, Hutchinson A, Crenshaw A, Cancel-Tassin G, Staats BJ, Wang Z, Gonzalez-Bosquet J, Fang J, Deng X, Berndt SI, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cussenot O, Valeri A, Andriole GL, Crawford ED, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover R, Hayes RB, Hunter DJ, Chanock SJ. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet. 2008; 40(3):310–315. [PubMed: 18264096]
14. Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? Bmj. 1999; 319(7224):1562–1565. [PubMed: 10591726]
15. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol. 2004; 159(9):882–890. [PubMed: 15105181]
16. Zheng SL, Sun J, Wiklund F, Smith S, Stattin P, Li G, Adami HO, Hsu FC, Zhu Y, Balter K, Kader AK, Turner AR, Liu W, Bleecker ER, Meyers DA, Duggan D, Carpten JD, Chang BL, Isaacs WB, Xu J, Gronberg H. Cumulative association of five genetic variants with prostate cancer. N Engl J Med. 2008; 358(9):910–919. [PubMed: 18199855]
17. Schmidt C. Personal genetic tests facing scrutiny. J Natl Cancer Inst. 2008; 100(6):382–383. 386. [PubMed: 18334700]
18. Stanford JL, Wicklund KG, McKnight B, Daling JR, Brawer MK. Vasectomy and risk of prostate cancer. Cancer Epidemiol Biomarkers Prev. 1999; 8(10):881–886. [PubMed: 10548316]
19. Agalliu I, Salinas CA, Hansten PD, Ostrander EA, Stanford JL. Statin use and risk of prostate cancer: results from a population-based epidemiological study. Am J Epidemiol. 2008 In press.
20. Sambrook, J.; Russell, DW. Molecular cloning : a laboratory manual. 3. Vol. 3. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press; 2001. p. v(various pagings) p
21. Harlow BL, Davis S. Two one-step methods for household screening and interviewing using random digit dialing. Am J Epidemiol. 1988; 127(4):857–863. [PubMed: 3354550]
22. Breslow NE, Day NE. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC Sci Publ. 1980; (32):5–338. [PubMed: 7216345]
23. Cole P, MacMahon B. Attributable risk percent in case-control studies. Br J Prev Soc Med. 1971; 25(4):242–244. [PubMed: 5160433]
24. Surveillance Epidemiology, and End Results (SEER) Program. National Cancer Institute, DCCPS, Surveillance Research Program, Statistical Research and Applications Branch; April. 2008 (www.seer.cancer.gov). SEER 9, 30 Yr Limited Duration Prevalence on 1/1/2005 by Duration
25. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological). 1972; 34(2):187–220.
26. Cornfield J. A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix. J Natl Cancer Inst. 1951; 11(6):1269–1275. [PubMed: 14861651]
27. Sinclair JC, Bracken MB. Clinically useful measures of effect in binary analyses of randomized trials. J Clin Epidemiol. 1994; 47(8):881–889. [PubMed: 7730891]
28. Zhong H, Prentice RL. Bias-reduced estimators and confidence intervals for odds ratios in genome-wide association studies. Biostatistics. 2008

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

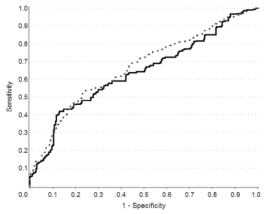

Figure 1.
ROC curves for the effect of five SNPs in the 8q24, 17q12, and 17q243 chromosomal regions in predicting risk of prostate cancer. Solid line indicates curve fit with clinical variables only (i.e., age, serum PSA level at diagnosis or interview, first-degree family history of prostate cancer) and dashed line indicates curve fit with clinical variables plus the five SNPs.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

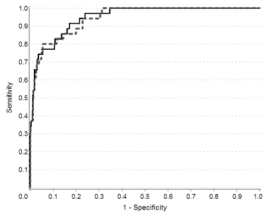

Figure 2.

ROC curves for the effect of five SNPs in the 8q24, 17q12, and 17q24.3 chromosomal regions in predicting ten-year prostate cancer-specific mortality. Solid line indicates curve fit with clinical variables only (i.e., age at diagnosis, stage, Gleason score, serum PSA level at diagnosis, first-degree family history of prostate cancer, and primary treatment) and dashed line indicates curve fit with clinical variables plus the five SNPs.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table I

Demographic and clinical characteristics of Caucasian study participants

Characteristic	Cases		Controls	
	n=1308	%	n=1266	%
Age at reference date, years				
35–49	102	7.8	107	8.5
50–54	188	14.4	178	14.1
55–59	325	24.9	343	27.1
60–64	395	30.2	334	26.4
65–69	153	11.7	160	12.6
70–74	145	11.1	144	11.4
1st Degree family history of PC				
No	1025	78.4	1125	88.9
Yes	283	21.6	141	11.1
PSA at diagnosis or interview (ng/mL)*				
0–3.9	178	13.6	351	27.7
4.0–9.9	721	55.1	33	2.6
10.0–19.9	190	14.5	6	0.5
≥20.0	118	9.0	0	-
Missing	101	7.7	876	69.2
Stage				
Local	1,015	78.1		
Region	253	19.5		
Distant	32	2.5		
Missing	8			
Gleason score				
2–4	66	5.1		
5–6	681	52.2		
7=3+4	355	27.2		
7=4+3	76	5.8		
8–10	126	9.7		
Missing	4			
Primary treatment				
Radical prostatectomy	770	58.9		
Radiation	352	26.9		
Androgen deprivation therapy	60	4.6		
Other treatment	11	0.8		
Active surveillance	115	8.8		

*Serum PSA level measured at diagnosis for cases and at interview for a random sample of controls from Study I.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table II

Risk of prostate cancer associated with SNPs in the 8q24, 17q12, and 17q24.3 chromosomal regions

Variable or SNP	Region	Associated risk group	Study I & II*						Study I only†					
			Frequency of associated factor (%)		OR	95% CI		P-value‡	Frequency of associated factor (%)		OR	95% CI		P-value‡
			Cases (n=1308)	Controls (n=1266)					Cases (n=475)	Controls (n=364)				
Family history of PC		Yes	20.9	10.8	2.32	1.85	2.92	4.3×10^{-29}	19.0	10.7	1.86	1.04	3.35	0.04
rs4430796	17q12	AA	32.9	25.0	1.43	1.19	1.71	9.9×10^{-6}	31.8	26.7	1.52	0.97	2.41	0.07
rs1859962	17q24.3	GG	27.3	23.2	1.25	1.03	1.51	6.6×10^{-4}	29.3	25.0	1.41	0.89	2.24	0.01
rs6983561	8q24	CC+CA	11.1	6.5	1.76	1.30	2.38	4.0×10^{-6}	11.8	3.9	2.63	1.17	6.10	0.02
rs6983267	8q24	GG+GT	80.8	75.1	1.34	1.10	1.64	6.7×10^{-4}	81.9	74.2	1.08	0.66	1.79	0.76
rs1447295	8q24	AA+AC	25.2	19.4	1.34	1.10	1.63	3.3×10^{-14}	25.3	19.5	1.35	0.82	2.22	0.24

*Adjusted for age at reference date; men with missing genotype information for any SNP were excluded from the analysis (n=97 cases, n=58 controls).

†Two-sided p-values were calculated by comparing the -2 Log Likelihood difference between nested models.

‡Models were adjusted for serum PSA in addition to age at reference date and included cases and controls from Study I with information for PSA.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table III

Population attributable risk percent (PAR%) for SNPs in the 8q24, 17q12, and 17q24.3 chromosomal regions

Variable or SNP	Associated risk group	Frequency of associated factor (%)		PAR%	Adjusted PAR%*	
		Cases	Controls		3% Prevalence	5% Prevalence
1st degree family history of PC	Yes	20.9	10.8	12.5	11.8	11.0
rs4430796	AA	32.9	25.0	9.7	9.4	9.1
rs1859962	GG	27.3	23.2	5.4	5.3	5.2
rs6983561	CC+CA	11.1	6.5	4.7	4.5	4.3
rs6983267	GG+GT	80.8	75.1	20.4	19.8	19.0
rs1447295	AA+AC	25.2	19.4	6.2	6.0	5.8
All five SNPs	All†	0.3	0.1	39.2	38.1	37.0
All five SNPs and family history	All‡	0.1	0	46.7	45.4	45.0

*PAR% adjusted for inflation due to estimating relative risk using the odds ratios, at the indicated probability of being diagnosed with prostate cancer.

†Men with five at-risk genotypes.

‡Men with five at-risk genotypes plus a positive family history of prostate cancer.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table IV

Risk of prostate cancer associated with cumulative effects of SNPs in the 8q24, 17q12, and 17q24.3 chromosomal regions and family history of prostate cancer (PC)

Variable	Cases		Controls		Odds ratio*	95% CI		P-value†,‡
	n=1308	%	n=1266	%				
Genotypes at five SNPs §								
1st degree family history of PC					2.31	1.84	2.91	9.5×10^{-12}
No. of associated genotypes								
0	80	6.6	143	11.8	1.00	Reference		
1	431	35.6	499	41.3	1.48	1.09	2.01	
2	450	37.2	421	34.9	1.88	1.38	2.56	
3	206	17.0	122	10.1	2.97	2.08	4.26	
≥4	44	3.6	23	1.9	3.36	1.90	6.08	1.2×10^{-12}
Genotypes at five SNPs and family history‖								
No. of associated factors								
0	65	5.4	130	10.8	1.00	Reference		
1	327	27.0	461	38.2	1.41	1.02	1.97	
2	481	39.7	425	35.2	2.25	1.63	3.13	
3	258	21.3	150	12.4	3.43	2.40	4.94	
4	70	5.8	38	3.2	3.65	2.24	6.03	
≥5	10	0.8	4	0.3	4.92	1.58	18.53	1.5×10^{-20}

*Adjusted for age at reference date; men with missing genotype information for any SNP were excluded from the analysis (n=97 cases, n=58 controls).

†Two-sided p-values were calculated by comparing the −2 Log Likelihood difference between nested models.

‡P-values for the number of associated genotypes and family history were calculated based on the Cochran-Armitage test for trend.

§Testing for the cumulative effect of five SNPs (rs4430796, rs1859962, rs6983561, rs6983267, rs1447295), adjusted for age and family history.

‖Testing for the cumulative effect of the five SNPs and family history of prostate cancer, adjusted for age.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table V

Prostate cancer-specific mortality associated with SNPs in the 8q24, 17q12 and 17q24.3 chromosomal regions

Variable or SNP (alleles)	Chromosomal region	Associated risk group*	Frequency of associated SNP or family history (%)		Hazard ratio		
			Cases† (n=1162)	PC-specific deaths (n=45)	HR‡	95% CI	
1st degree family history of PC		Yes	21.8	11.1	0.29	0.07	1.20
rs4430796 (G,A)	17q12	AA	32.2	40.5	1.34	0.65	2.77
rs1859962 (T,G)	17q24.3	GG	28.0	23.3	0.40	0.15	1.04
rs6983561 (A,C)	8q24	CC+CA	11.4	11.9	1.86	0.68	5.10
rs6983267 (T,G)	8q24	GG+GT	80.7	81.0	0.86	0.34	2.17
rs1447295 (C,A)	8q24	AA+AC	25.0	23.8	0.88	0.40	1.93

*Groups associated with elevated risk of prostate cancer in Table II.

†Living cases and those who died of other causes.

‡Cox proportional hazards model was used to estimate hazard ratios associated with each factor. The model was adjusted for age, Gleason score (2–6, 7=3+4, 7=4+3, 8–10), stage at diagnosis (local, regional, distant), PSA at diagnosis, and primary treatment.