Validation of Genome-Wide Prostate Cancer Associations in Men of African Descent

Bao-Li Chang[1], **Elaine Spangler**[1], **Stephen Gallagher**[1], **Christopher A. Haiman**[3], **Brian Henderson**[3], **William Isaacs**[4], **Marnita L. Benford**[5], **LaCreis R. Kidd**[5], **Kathleen Cooney**[6], **Sara Strom**[7], **Sue Ann Ingles**[8], **Mariana C. Stern**[8], **Roman Corral**[8], **Amit D. Joshi**[8], **Jianfeng Xu**[9], **Veda N. Giri**[10], **Benjamin Rybicki**[11], **Christine Neslund-Dudas**[11], **Adam S. Kibel**[12], **Ian M. Thompson**[13], **Robin J. Leach**[13], **Elaine A. Ostrander**[14], **Janet L. Stanford**[15], **John Witte**[16], **Graham Casey**[17], **Rosalind Eeles**[18] **for The UKGPCS Coordinating Group and Collaborators**, **Ann W. Hsing**[19], **Stephen Chanock**[19], **Jennifer J. Hu**[20], **Esther M. John**[21], **Jong Park**[22], **Klara Stefflova**[1], **Charnita Zeigler-Johnson**[1], and **Timothy R. Rebbeck**[1,2]

[1]Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia

[2]Abramson Cancer Center, University of Pennsylvania School of Medicine, Philadelphia

[3]Department of Preventive Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles

[4]Johns Hopkins University School of Medicine and Bloomberg School of Public Health, Baltimore, MD

[5] University of Louisville, KY

[6]Department of Medicine, University of Michigan Medical School, Ann Arbor, MI

[7]MD Anderson Cancer Center, Houston, TX

[8]Department of Preventive Medicine, University of Southern California, Keck School of Medicine, Los Angeles CA

[9]Wake Forest University, Winston-Salem, NC

[10] Fox Chase Cancer Center, Philadelphia, PA

[11] Henry Ford Hospital, Detroit, MI

[12]Department of Surgery and Genetics, Washington University School of Medicine, St. Louis, MO

[13]Department of Urology and the Cancer Therapy and Research Center, University of Texas Health Science Center at San Antonio, TX

[14] National Human Genome Research Institute, Bethesda, MD

[15] Fred Hutchinson Cancer Research Center, Seattle, WA

[16] Departments of Epidemiology & Biostatistics, Urology, Institute for Human Genetics, University of California, San Francisco, CA

[17] Department of Preventive Medicine, Keck School of Medicine, University of Southern California

Corresponding Author: Timothy R. Rebbeck, Ph.D. Department of Biostatistics and Epidemiology University of Pennsylvania School of Medicine 217 Blockley Hall, 423 Guardian Drive Philadelphia, PA 19104-6021 rebbeck@mail.med.upenn.edu Telephone: 215-898-1793, Fax: 215-573-1050.

[18]Institute for Cancer Research and Royal Marsden NHS Trust, Sutton, UK

[19] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD

[20] University of Miami Miller School of Medicine, Miami, FL

[21]Northern California Cancer Center, Union City, CA

[22]Moffitt Cancer Center, Tampa, FL

Abstract

Background—Genome-wide association studies (GWAS) have identified numerous prostate cancer susceptibility alleles, but these loci have been identified primarily in men of European descent. There is limited information about the role of these loci in men of African descent.

Methods—We identified 7,788 prostate cancer cases and controls with genotype data for 47 GWAS-identified loci.

Results—We identified significant associations for SNP rs10486567 at *JAZF1*, rs10993994 at *MSMB*, rs12418451 and rs7931342 at 11q13, and rs5945572 and rs5945619 at *NUDT10/11*. These associations were in the same direction and of similar magnitude as those reported in men of European descent. Significance was attained at all report prostate cancer susceptibility regions at chromosome *8q24*, including associations reaching genome-wide significance in region 2.

Conclusion—We have validated in men of African descent the associations at some, but not all, prostate cancer susceptibility loci originally identified in European descent populations. This may be due to heterogeneity in genetic etiology or in the pattern of genetic variation across populations.

Impact—The genetic etiology of prostate cancer in men of African descent differs from that of men of European descent.

Keywords

prostate cancer; genetic susceptibility; men of African descent

Introduction

The differences in prostate cancer incidence and mortality across men of different racial groups are well documented. According to SEER, prostate cancer has a age-adjusted incidence rate of 234.6 per 100,000 in African American and 150.4 per 100,000 in European American men. Additionally, a 2.4-fold difference in mortality rate (62.3 per 100,000 in African Americans vs. 25.6 per 100,000 in European Americans) represents the greatest disparity between these groups of any major cancer site. Despite this profound public health concern, knowledge of the etiological underpinnings for this disparity remains unclear. It is likely that inherited susceptibility, environmental exposures, lifestyle, behavior, screening, and cancer treatment all influence the disparity between men of different racial and ethnic backgrounds.

A number of recent genome-wide association studies (GWAS) have identified numerous prostate cancer susceptibility loci including *CTBP2* (chr. 10q26), *EHPB1* (chr. 2p15) , *HNF1B* (chr. 17q12), *IGF2/IGF2A/INS* (chr. 11p15), *ITGA6* (chr. 2p31), *KLK2/3* (chr. 19q13), *LMTK2* (chr. 7q21), *MSMB* (chr. 10q11), *NKX3.1* (chr. 8p21), *NUDT10/11* (chr. Xp11.22), *PDLIM5* (chr. 4q22), *SELB* (chr. 3q21.3), *SLC22A3* (chr. 6q25), *TET2* (chr. 4q24), *THADA* (chr. 2p21), *TTLL1/BIK/MCAT/PACSIN2* (chr. 22q13), as well as loci on chromosome 11q13, 17q12, 17q24, and multiple regions at chromosome 8q24 (1-17). These

loci were discovered primarily in European descent men (EDM), with the exception being the prostate cancer susceptibility loci at chromosome 8q24, which were identified by linkage and admixture mapping (15, 18). Studies suggest that some genetic variants confer risk across populations but with different magnitudes of the risk in different populations, or they may only confer risk in one population but not in others [11,19]. Because the prevalence of prostate cancer and the allele frequencies differ between EDM and African descent men (ADM), it is important to estimate the effects of these GWAS risk variants originally identified in EDM on ADM before generalization of the GWAS associations in ADM. Three recent studies have also attempted to validate associations between some of the loci listed above and prostate cancer in ADM. Xu et al. [36] studied 868 cases and 878 controls and validated the loci at 8q24 ($p=0.034$ to $p=2\times10^{-5}$) and 3p12 ($p=0.029$). Waters et al. (19) studied 860 cases and 575 controls, and validated *KLK2/3* (19q13.33) and *NUDT10/11* (Xp11.22). Finally, Hooker et al. [51] validated 8q24 ($p=1x10^{-4}$), 11q13.2 ($p=0.009$), *HNF1B/TCF2* (17q12; $p=0.008$), *KLK2/3* (19q13.33; $p=0.04$), and *NUDT11* (Xp11.22; $p=0.05$) in 454 cases and 301 controls. The validated loci were not consistent across these studies, perhaps due to relatively small sample sizes in each study. To confirm associations at previously identified prostate cancer susceptibility loci in ADM, we obtained data from 7,788 ADM from 19 centers in the US and the UK for pooled analyses of GWAS-identified loci and prostate cancer.

Methods

Study Sample

The sample studied here consisted of 4,040 cases and 3,748 controls ascertained from 19 centers (Supplementary Table 1). A detailed description of each center's study is presented in Appendix 1 and a summary of the study methods is presented in Supplementary Table 5. These studies include the Prostate Cancer Genetics Studies (CaP Genes) at the University of California, San Francisco (20), Fred Hutchinson Cancer Research Center (FHCRC) Prostate Cancer Studies (21, 22), The Prostate Risk Assessment Program (PRAP) at Fox Chase Cancer Center (23), The Flint Men's Health Study (FMHS) (24, 25), Gene-Environment Interaction in Prostate Cancer (GECAP) Study at Henry Ford Hospital (26), Los Angeles County Study (LACS) (27), Prostate Cancer Clinical Outcome Study (PC^2OS) at the University of Louisville(28), MD Anderson Cancer Center (29), The Multiethnic Cohort Study (MEC) (30), Moffitt Cancer Center Study (31), NCI Prostate Tissue Study (NCIPTS), University of Pennsylvania Study of Cancer Outcomes, Risk, and Ethnicity (SCORE) (32), University of Texas San Antonio Center for Biomarkers of Risk for Prostate Cancer (SABOR), University of Texas Health Science Center at San Antonio (33, 34), San Francisco Bay Area Prostate Cancer Study (SFBAPCS) (35), United Kingdom Genetic Prostate Cancer Study (UKGPCS), Wake University Consortium including participants from the Johns Hopkins University, Wake Forest University, and Washington University (36). Two of these studies, SFBAPCS and UKGPCS, have only contributed to case-case analyses of disease aggressiveness since only cases were available from these two studies. SNPs were chosen if they were implicated in previous GWAS studies (1-3, 37), in follow-up fine-mapping studies (5-7, 38, 39), or associated with disease aggressiveness (4, 40). Available SNPs in all regions of 8q24, some of which were initially identified through linkage and admixture mapping in ADM and confirmed in GWAS studies, were also included (10, 11, 14-16, 41).

Genotype data were excluded if they were found to have genotyping failure rates >5% within each study center or if they deviated significantly from Hardy-Weinberg proportions. We set a threshold of $p<0.001$ based on multiple-test adjustment for the number of SNPs tested (family-wise error rate $p=0.05$ divided by 50 SNPs equals to $p=0.001$). SNPs were included in the present analysis if we obtained at least 1000 genotypes in cases and controls

from the contributing centers by October 2009. A summary of the data contributed by each center by SNP is summarized in Supplementary Table 6.

Statistical Methods

Departure from Hardy-Weinberg equilibrium was assessed for each SNP in control subjects of the combined study populations using the Chi-Square Goodness-of-Fit Test. Any SNP that showed departure from Hardy-Weinberg equilibrium with $p < 0.001$ in controls was excluded from subsequent analyses. Unconditional logistic regression models were used to estimate odds ratios (ORs) and 95% confidence intervals (95% CIs) to measure the association between individual SNP genotypes and prostate cancer risk or disease aggressiveness defined as Gleason score <7 vs. 7+ or tumor stage T1/T2 vs. T3/T4. Analyses were undertaken using an additive mode of inheritance, adjusting for age and study centers (results shown in Table 1 and Supplementary Table 2-4).

Subgroup analyses were also carried out in order to estimate whether African ancestry affected the reported associations. This analysis included a subset of study centers for which estimated percentage (%) of African ancestry was available (Supplemental Table 5). Centers used different ancestry informative marker (AIMs) panels (Supplementary Table 5). These AIMs were obtained from the original genotyping methods used in each center, and were comparable based on several measures of marker informativeness (FST, FIC, and δ). The statistical methods used to estimate ancestry proportion, STRUCTURE and ANCESTRYMAP, have used same hierarchical model and probabilistic measures and would results in similar/high correlated measurements. In addition, we analyzed data stratifying by center to adjust for potential confounding by ancestry proportion within each participating study and to minimize the influence of varying informativeness of AIM panels.

These studies include nested case-control studies from within cohorts, matched and unmatched case-control studies, as well as case-only series. To address the potential study heterogeneity, age-adjusted ORs and 95% CIs for SNPs were estimated for each study population separately, and forest plots were generated for independent SNPs with p-values < 0.05 (Supplementary Figure 1). Potential heterogeneity in the association of SNPs with prostate cancer among study populations was examined by Breslow-Day homogeneity test. All statistical analyses were performed using SAS 9.2 and PLINK(42). An LD heat map (Figure 1) was generated based on HapMap YRI data using Haploview(43). Inferences were made using two-sided hypothesis testing with a p-value <0.05. Because this is a validation study, we did not correct for multiple hypothesis tests.

Results

We were able to validate some, but not all, prostate cancer GWAS loci (Table 1 for SNPs outside of 8q24 regions, and Table 2 for SNPs located within 8q24). Most associations reported here were in the same direction and with an equal or smaller magnitude as those originally reported in EDM. However, a number of associations reported here were not in the same direction as those reported in EDM (i.e., CTBP2, 11q13, and 22q13, Table 1), suggesting these alleles are not consistent with prostate cancer risk in ADM. A number of loci that were implicated in EDM were not associated with prostate cancer risk in ADM. These included *CTBP2* (rs4962416), 11q13 (rs12418451), *IL16* (rs4072111), *CDH13* (rs4782726), and 22q13 (rs9623117) with OR<1 (i.e., in the opposite direction from that reported in EDM), and *EHBP1* (rs721048), *LMTK2* (rs6465657), *MINPP1* (rs12771728), Chromosome 12 (rs902774), and *KLK2/3* (rs887391) with OR near 1.0. Furthermore, the upper bound of the 95% confidence interval for a number of loci in ADM did not overlap at least earlier estimates made in EDM, including 3p12.1 (rs2660753), *DAB2IP* (rs1571801), *MSMB* (rs10993994), *CTBP2* (rs4962416), *HNF1B* (rs4430796 and rs7501939), *KLK2/3*

(rs2735839), 22q13 (rs9623117), and *NUDT10/11* (rs5945572 and rs5945619). These results suggest that some loci with genome-wide significance in non-African descent populations may not be associated with prostate cancer or may not have the same magnitude of effect in ADM.

Several SNPs showed statistically significant associations. SNPs in *JAZF1* (rs10486567; OR=1.18, p=0.0002), *MSMB* (rs10993994; OR=1.12, p=0.005), 11q13 (rs10896449, OR=1.12, p=0.031 and rs7931342, OR=1.15, p=0.014) and *NUDT10/11* (rs5945572, OR=1.11, p=0.02 and rs5945619, OR=1.09, p=0.039) were statistically significantly associated with prostate cancer risk. . The direction of effect of each of these associations was in the same direction as those reported in EDM (Table 1).

We also undertook a similar analysis that excluded data that have been published previously to isolate a subset of study centers for evaluating further evidence of independent replications (19, 36). After excluding data from those studies (i.e., JHU, MEC, Wake-Hu, Wake-NC, and Wash U), both *JAZF1* rs10486567 (p=0.005) and *MSMB* rs10993994 (p=0.009) remained statistically significant. In both cases, the OR estimate in the subset trended away from the null hypothesis (OR=1.23 in the subset vs. 1.18 in the total sample, and OR=1.17 in the subset vs. 1.12 for the total sample, respectively). SNP rs10896449 at 11q13 stayed nominally significant (p=0.02), but SNPs at *NUDT10/11* and SNP rs7931342 at 11q13 were no longer significant. These results further provide support for the association of *JAZF1* and *MSMB* with prostate cancer risk in ADM. While we were unable to mutually adjust for the effects of multiple SNPs in a single locus for the majority of loci, after mutually adjusting for multiple SNPs at 11q13, both SNP rs7931342 (OR=1.0, 94%CI: 0.77-1.30, p=0.999) and rs10896449 (OR=1.18, 95% CI: 0.93-1.49, p=0.17) became non-significant. Since the sample size for this last analysis is smaller than for the overall sample (i.e., N=2,013 vs. N= 3,954 or 4,463), we were not able to unambiguously determine which SNP contributed independently to the association signal seen at this locus. After mutual adjustment, the point estimates for rs7931342 changed from 1.15 to 1.0; rs10896449 changed from 1.12 to 1.18. These results suggest that rs10896449 or other SNPs in tight LD with rs10896449 maybe the SNP that contributes to the association signal at 11q13 locus. Multiple independent loci on chromosome 8q24 have been identified as playing a role in prostate cancer etiology. We were able to validate the association of each of these regions at 8q24 (Figure 1 and Supplementary Table 2). We had statistically significant evidence at the genome-wide association level for associations with regions 2 (rs13254738, rs6983561, and rs16901979), and statistically significant associations in region 1 (rs10090154), region 3 (rs6983267 and rs7000448), region 4 (rs7008482) and the region centromeric to region 2 (rs10086908). We also removed data that had been included in previous studies (19, 36, 44) of loci at 8q24. Significant associations remained for Regions 2 (block 2), 3 (block 4), 4 (centromeric to block 1) and the region centromeric to Region 2 (block 1). However, the marginal associations in region 1 (block 5) were no longer significant after the data from the published reports were excluded.

Because we have studied an admixed population of ADM, we also investigated potential bias due to population stratification by comparing the association results with or without adjusting for percentage of non-African ancestry estimated from ancestry informative markers (AIMs). Ancestry adjustment analyses were undertaken in 8 of the 19 centers for which AIMs data were available (Supplementary Table 3). We observed significant differences in the proportion of African ancestry across centers ($\chi^2_{7\text{-Kruskal-Wallis}}=339.6$, p<0.0001). However, these differences may reflect not only known geographic differences in ADM admixture (45), but also the different ancestry marker panels and methods used to estimate the ancestry proportions across centers (Supplementary Table 5). Therefore, we have performed all analyses with adjustment for center effects to reduce the impact of

different ancestry marker panels and methods used across centers. Among those centers with ancestry marker data, inclusion of percent non-African ancestry did not substantially change the associations or inferences for any locus compared with models adjusted only for age and center.

We also evaluated the effect of the GWAS SNPs studied here on prostate cancer aggressiveness by repeating the analysis with stratification by clinical (TNM) stage and histologic (Gleason) grade (Supplementary Table 4). For SNPs that showed a significant association in the comparisons of both high grade/stage against controls and low grade/stage against controls, there were no statistically significant differences between high and low grade/stage cases. A number of loci were associated with disease aggressiveness, but in no instance was there evidence for statistically significant differences in the associations by disease aggressiveness after correction for multiple testing (Supplementary Table 4).

We also evaluated whether there was evidence for first-order interactions between any of the loci identified as having a statistically significant main effect on risk of prostate cancer (Table 1). Using an additive (per-allele) model adjusted for age and study center, we considered interactions only among SNPs not in LD. The most significant interaction identified was between two SNPs on chromosome 8q24: rs10086908 (centromeric to Region 2) and rs6983267 (Region 3; nominal p-value=0.021). However, after correction for multiple testing using the FDR, this interaction was no longer significant (FDR p-value=0.42). No other p-values for interaction reached statistical significance.

Finally, we evaluated whether there was evidence for heterogeneity in associations across centers by generating forest plots of the individual center OR estimates that reached overall statistical significance (Supplementary Figure 1). With very few exceptions, the associations that reached any level of significance showed remarkable consistency in the direction of the risk estimates. There was no statistically significant heterogeneity in effects across centers (p>0.05 for all SNPs).

Discussion

A number of recent reports have modeled the role of genomic markers on prostate cancer susceptibility (1-9). We have validated a number of these loci, including 8q24, *JAZF1*, *MSMB*, 11q13, and *NUDT10/11*. In general, the point estimates of risk at these loci in our current pooled analysis of 19 studies suggest that the effects of these loci in ADM are similar to those in EDM. We also observed no statistically significant heterogeneity of effects across studies (Supplementary Figure 1). A number of loci were not validated in our analysis, despite reaching genome-wide significance in GWAS studies of EDM. This discrepancy may be explained in a number of ways. First, the present study may not have been powered to identify very small effects of these loci. However, for a number of loci, we estimated ORs<1.0 with 95% confidence intervals that do no overlap the OR estimates originally reported in EDM. The effects of most remaining non-significant associations were obtained with OR<1.05, which are lower than those estimated in EDM. If the effects of these alleles are in fact smaller in magnitude in ADM than those reported in EDM, the present study may not have been able to detect these effects. Second, allele frequencies in EDM and ADM differ at many of the loci studied here (Table 1), as do patterns of linkage disequilibrium by ethnicity (46). These differences also may affect the ability to detect significant effects at some loci in ADM, where they may have been detectable in EDM. However, the reverse situation is also possible (Table 1). Finally, if none of these limitations applies, it is possible that the loci not validated in the present study confer susceptibility only in EDM, but not ADM. While it is unlikely that there are substantial biological differences in prostate cancer etiology between EDM and ADM, interactions of

environmental exposures, prostate cancer screening, and other non-genetic risk factors may influence the penetrance of these alleles that may manifest in different risk profiles.

One of the more consistent associations identified to date is that of rs10993994 at *MSMB* (10q11) (2, 3), which is confirmed as a prostate cancer susceptibility locus in ADM in this study. *MSMB* is a microseminoprotein beta gene that encodes PSP94, a nonglycosylated cysteine-rich protein that is a member of the immunoglobulin binding factor family synthesized by epithelial cells in the prostate and secreted into seminal plasma (3). While the exact function of PSP94 is not well-established, it is postulated to be involved in growth regulation, gene expression, and apoptosis in prostate cancer cells (2). PSP94 and its binding protein in serum, PSPBP, are potential serum markers for both prostate cancer risk and aggressiveness (47, 48), unlike the current PSA screening which mainly detects the presence of prostate cancer (47). The effect of rs10993994 in *MSMB* gene expression has been investigated in function studies (5, 39). The prostate cancer risk associated T allele of the rs10993994 SNP had only 13% of the promoter activity compared with the C allele, and treatment with increasing concentrations of the synthetic androgen R1881 resulted in a dose-dependent increase in promoter activity of the C, but not the T allele of the this SNP. Additionally, tumor cell lines with a CC or CT genotype revealed a high level of *MSMB* gene expression compared with cell lines with a TT genotype. These findings were specific to the alleles of rs10993994 and not from other SNPs in the proximal promoter of *MSMB*. The significant association found in rs10993994 and lack of association found in two other *MSMB* SNPs included in our study also suggests the potential of rs10993994 as the causal SNP. Further fine-mapping studies that take advantage of the shorter LD pattern in ADM would serve to augment this hypothesis.

JAZF1 ("juxtaposed with another zinc finger protein 1") was identified by the Cancer Genetic Markers of Susceptibility (CGEMS) study as associated with prostate cancer case-control status (3). This same group has undertaken fine mapping at this locus and confirmed that the original GWAS association with rs10486567 (the SNP validated in ADM here) is likely to be the marker responsible for the association signal at this locus (49). Since rs10486567 lies in intron 2 of *JAZF1* and is not known to alter any apparent splicing or expression of this gene, the functional significance of this association has yet to be determined. JAZF1 has been associated with somatic fusion proteins in endometrial tumors (50-53), but no other genomic associations have been reported.

Two previous studies (19, 54) suggested that *NUDT10/11* was associated with prostate cancer in ADM. One study of ADM, not included in the present data, also reported that SNPs at 11q13 were associated with prostate cancer in ADM(54). The marginal association between these two loci and prostate cancer in this study is suggestive of validation with GWAS associations in EA populations, but additional data may be required to fully validate these associations in ADM.

We have also validated the previously reported associations of multiple regions of chromosome 8q24 and prostate cancer in ADM. Originally identified by admixture mapping methods, and verified in GWAS (18), this locus has been shown to be comprised of a number of independent prostate cancer susceptibility regions (11, 41, 55, 56). Multiple regions have been validated in our study, with the strongest association signals seen in regions 2 and 3, and our findings are consistent with the fine-mapping of the admixture scan [33]. The association signals seen in regions 1, 4, and a region centromeric to region 2 are much weaker compared to those in regions 2 and 3.

Finally, a number of other loci did not reach statistical significance in any analysis, and in fact provided no evidence for association with prostate cancer in ADM. These included

many loci that reached genome-wide levels of significance in EA but had p-value>0.2 (and many with p>0.9) in ADM (Table 1). These include associations that were reported by two studies of ADM that are included in the present analysis, but did not reach statistical significance in the current combined data set, including *KLK2/3*, and *HNF1B/TCF2* (19, 36).

It is possible that a number of these statistically non-significant associations were underpowered in the present sample, especially those based on loci with lower minor allele frequencies. However, the adjusted OR estimates in ADM were often substantially lower than those reported in EA men (Table 1). Indeed, some risk estimates in ADM that had been estimated to be OR>1 were estimated in ADM to be OR<1, suggesting no evidence for a comparable association in between the two groups. There are a number of possible explanations for these findings. First, the loci identified in GWAS studies of EDM populations could represent false positive associations that cannot be replicated in ADM. Given the large sample sizes in replication studies and strong p-values associated with these loci in previous reports, this is an unlikely scenario. Second, there may be real heterogeneity in prostate cancer etiology that may be reflected by differences in allele frequency (i.e., ability to detect associations) or differences in the context in which these alleles are acting in EDM vs. ADM due to differences in environmental exposures, lifestyle or other effect modifiers not measured in studies to date. The present data do not allow us to address whether prostate cancer in ADM is less strongly influenced by genes relative to other factors than in EDM. However, the present results should be considered in future studies that may attempt to address this hypothesis. Third, the causal variants may not have been identified and genotyped yet, and the causal variants may be different in EDM and ADM. This question cannot be resolved by the data presented here, and will require additional fine-mapping studies as well as ADM-specific GWAS studies in which existing GWAS loci may be validated and new loci may be identified.

Despite the validation of some prostate cancer loci in ADM, there was no strong evidence that these loci had different effects on advanced (e.g., high stage or grade) disease compared with less advanced disease (e.g., low stage or grade). This may in part be due to limited power to detect significant differences between men with more vs. less aggressive disease features. In some cases, there were suggestions that some SNPs were associated with more aggressive disease, including a number of SNPs at Chr. 8q24 (rs6981122, rs7000448, rs16901896) as well as others such as rs7904463 (Chr. 10) and rs5945572 (Chr. X). In these cases, there is a suggestion of stronger associations in more vs. less aggressive disease in a case-control study design, but there were no statistically significant differences observed between more and less aggressive cases in a case-case comparison. Similarly, there were a number of loci for which the association was stronger for less advanced disease compared with more advanced disease. These included the associations for rs9623117 at 22q13, *MSMB* and *JAZF1* SNPs, for which the overall significant association among all cases combined (Table 1) appeared to exist only in cases with less aggressive features (Supplementary Table 4). Our results in ADM are in consistent with the report by Kader et al. (57) that showed the majority of currently identified GWAS risk-associated SNPs could not differentiate aggressive from less aggressive diseases in EDM. However, contrary to the significant finding in this report showing that SNPs in *KLK2/3* and *MSMB*, both related to serum PSA levels, were associated with less aggressive disease, our null finding in *KLK2/3* and *MSMB* implies that PSA screening may not introduce the same degree of bias in cancer detection in ADM as seen in EDM.

In studying an admixed population of ADM men, there is a concern for potential bias due to confounding by ethnicity (i.e., population stratification). To address the potential that there is bias in the risk estimates, we undertook a subset analysis of those centers that had

genotyped ancestry markers and estimated the proportion of African ancestry. We observed no substantial bias in the estimates of association for any SNP. In fact, compared to associations adjusted only for age and center, the odds ratios for 7 of 47 (15%) of associations adjusted for age, center, and percent non-African ancestry changed by 5% or more: three of these estimates moved away from the null hypothesis while 4 of these estimates changed toward the null . These empirical data suggest that the potential for bias due to population stratification is not large, and that the direction of this bias may not always be away from the null hypothesis. None of these SNPs was significantly associated with probability of having prostate cancer before or after adjustment for ancestry, so the consideration of ancestry did not change any inferences based on our results. Limitations of the approach used here include the use of different sets of markers and approaches to estimating African ancestry in only a subset of the available studies. However, our data provide no evidence for substantial bias due to population stratification in associations of GWAS SNPs in prostate cancer etiology.

In conclusion, we have validated in ADM the associations of some, but not all, prostate cancer susceptibility loci originally identified in non-African descent populations. The finding that the genetic etiology of prostate cancer may be different in ADM and EDM suggests that studies that take advantage of the shorter LD blocks in ADM or more complete resequencing efforts will facilitate identification of causal variants in verified risk loci.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Gudmundsson J, Sulem P, Rafnar T, et al. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet. 2008; 40:281–3. [PubMed: 18264098]

2. Eeles RA, Kote-Jarai Z, Giles GG, et al. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet. 2008; 40:316–21. [PubMed: 18264097]

3. Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet. 2008; 40:310–5. [PubMed: 18264096]

4. Duggan D, Zheng SL, Knowlton M, et al. Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. J Natl Cancer Inst. 2007; 99:1836–44. [PubMed: 18073375]

5. Lou H, Yeager M, Li H, et al. Fine mapping and functional analysis of a common variant in MSMB on chromosome 10q11.2 associated with prostate cancer susceptibility. Proc Natl Acad Sci U S A. 2009; 106:7933–8. [PubMed: 19383797]

6. Zheng SL, Stevens VL, Wiklund F, et al. Two independent prostate cancer riskassociated Loci at 11q13. Cancer Epidemiol Biomarkers Prev. 2009; 18:1815–20. [PubMed: 19505914]

7. Sun J, Zheng SL, Wiklund F, et al. Evidence for two independent prostate cancer riskassociated loci in the HNF1B gene at 17q12. Nat Genet. 2008; 40:1153–5. [PubMed: 18758462]

8. Gudmundsson J, Sulem P, Steinthorsdottir V, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet. 2007; 39:977–83. [PubMed: 17603485]

9. Hsu FC, Sun J, Wiklund F, et al. A Novel Prostate Cancer Susceptibility Locus at 19q13. Cancer Res. 2009

10. Al Olama AA, Kote-Jarai Z, Giles GG, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. Nat Genet. 2009; 41:1058–60. [PubMed: 19767752]

11. Haiman CA, Patterson N, Freedman ML, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet. 2007; 39:638–44. [PubMed: 17401364]

12. Eeles RA, Kote-Jarai Z, Al Olama AA, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet. 2009; 41:1116–21. [PubMed: 19767753]

13. Gudmundsson J, Sulem P, Gudbjartsson DF, et al. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. Nat Genet. 2009; 41:1122–6. [PubMed: 19767754]

14. Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet. 2007; 39:645–9. [PubMed: 17401363]

15. Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet. 2006; 38:652–8. [PubMed: 16682969]

16. Yeager M, Chatterjee N, Ciampa J, et al. Identification of a new prostate cancer susceptibility locus on chromosome 8q24. Nat Genet. 2009; 41:1055–7. [PubMed: 19767755]

17. Kote-Jarai Z, Easton DF, Stanford JL, et al. Multiple novel prostate cancer predisposition loci confirmed by an international study: the PRACTICAL Consortium. Cancer Epidemiol Biomarkers Prev. 2008; 17:2052–61. [PubMed: 18708398]

18. Freedman ML, Haiman CA, Patterson N, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proc Natl Acad Sci U S A. 2006; 103:14068–73. [PubMed: 16945910]

19. Waters KM, Le Marchand L, Kolonel LN, et al. Generalizability of associations from prostate cancer genome-wide association studies in multiple populations. Cancer Epidemiol Biomarkers Prev. 2009; 18:1285–9. [PubMed: 19318432]

20. Liu X, Plummer SJ, Nock NL, Casey G, Witte JS. Nonsteroidal antiinflammatory drugs and decreased risk of advanced prostate cancer: modification by lymphotoxin alpha. Am J Epidemiol. 2006; 164:984–9. [PubMed: 16931544]

21. Stanford JL, Wicklund KG, McKnight B, Daling JR, Brawer MK. Vasectomy and risk of prostate cancer. Cancer Epidemiol Biomarkers Prev. 1999; 8:881–6. [PubMed: 10548316]

22. Agalliu I, Salinas CA, Hansten PD, Ostrander EA, Stanford JL. Statin use and risk of prostate cancer: results from a population-based epidemiologic study. Am J Epidemiol. 2008; 168:250–60. [PubMed: 18556686]

23. Giri VN, Beebe-Dimmer J, Buyyounouski M, et al. Prostate cancer risk assessment program: a 10-year update of cancer detection. J Urol. 2007; 178:1920–4. discussion 4. [PubMed: 17868726]

24. Cooney KA, Strawderman MS, Wojno KJ, et al. Age-specific distribution of serum prostate-specific antigen in a community-based study of African-American men. Urology. 2001; 57:91–6. [PubMed: 11164150]

25. Heeringa SG, Alcser KH, Doerr K, et al. Potential selection bias in a community-based study of PSA levels in African-American men. J Clin Epidemiol. 2001; 54:142–8. [PubMed: 11166529]

26. Rybicki BA, Neslund-Dudas C, Nock NL, et al. Prostate cancer risk from occupational exposure to polycyclic aromatic hydrocarbons interacting with the GSTP1 Ile105Val polymorphism. Cancer Detect Prev. 2006; 30:412–22. [PubMed: 17067754]

27. Schwartz G, John E, Rowland G, Ingles S, Schwartz GG, John EM, Rowland G, Ingles SA. Prostate cancer in African American men and polymorphism in the calcium sensing receptor. Cancer Biology & Therapy. 2010 in press. Schwartz GG, John EM, Rowland G Ingles SA Prostate cancer in African American men and polymorphism in the calcium sensing receptor Cancer Biology & Therapy (in press). In Press.

28. Benford ML, VanCleave TT, Lavender NA, Kittles RA, Kidd LR. 8q24 sequence variants in relation to prostate cancer risk among men of African descent: a case-control study. BMC Cancer. 10:334. [PubMed: 20584312]

29. Strom SS, Yamamura Y, Flores-Sandoval FN, Pettaway CA, Lopez DS. Prostate cancer in Mexican-Americans: identification of risk factors. Prostate. 2008; 68:563–70. [PubMed: 18247399]

30. Kolonel LN, Henderson BE, Hankin JH, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol. 2000; 151:346–57. [PubMed: 10695593]

31. Park JY, Tanner JP, Sellers TA, et al. Association between polymorphisms in HSD3B1 and UGT2B17 and prostate cancer risk. Urology. 2007; 70:374–9. [PubMed: 17826523]

32. Rebbeck TR, Jaffe JM, Walker AH, Wein AJ, Malkowicz SB. Modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. J Natl Cancer Inst. Jun 16.1999 91(12): 1082. [see comment][erratum appears. Journal of the National Cancer Institute. 1998; 90:1225–9. [PubMed: 9719084]

33. Beuten J, Gelfond JA, Franke JL, et al. Single and multigenic analysis of the association between variants in 12 steroid hormone metabolism genes and risk of prostate cancer. Cancer Epidemiol Biomarkers Prev. 2009; 18:1869–80. [PubMed: 19505920]

34. Beuten J, Gelfond JA, Martinez-Fierro ML, et al. Association of chromosome 8q variants with prostate cancer risk in Caucasian and Hispanic men. Carcinogenesis. 2009; 30:1372–9. [PubMed: 19528667]

35. John EM, Schwartz GG, Koo J, Van Den Berg D, Ingles SA. Sun exposure, vitamin D receptor gene polymorphisms, and risk of advanced prostate cancer. Cancer Res. 2005; 65:5470–9. [PubMed: 15958597]

36. Xu J, Kibel AS, Hu JJ, et al. Prostate cancer risk associated loci in African Americans. Cancer Epidemiol Biomarkers Prev. 2009; 18:2145–9. [PubMed: 19549807]

37. Gudmundsson J, Sulem P, Manolescu A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet. 2007; 39:631–7. [PubMed: 17401366]

38. Hsu FC, Sun J, Wiklund F, et al. A novel prostate cancer susceptibility locus at 19q13. Cancer Res. 2009; 69:2720–3. [PubMed: 19318570]

39. Chang BL, Cramer SD, Wiklund F, et al. Fine mapping association study and functional analysis implicate a SNP in MSMB at 10q11 as a causal variant for prostate cancer risk. Hum Mol Genet. 2009; 18:1368–75. [PubMed: 19153072]

40. Sun J, Zheng SL, Wiklund F, et al. Sequence variants at 22q13 are associated with prostate cancer risk. Cancer Res. 2009; 69:10–5. [PubMed: 19117981]

41. Zheng SL, Sun J, Cheng Y, et al. Association between two unlinked loci at 8q24 and prostate cancer risk among European Americans. J Natl Cancer Inst. 2007; 99:1525–33. [PubMed: 17925536]

42. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–75. [PubMed: 17701901]

43. Haploview. [cited; Available from: http://www.broad.mit.edu/mpg/haploview

44. Salinas CA, Kwon E, Carlson CS, et al. Multiple independent genetic variants in the 8q24 region are associated with prostate cancer risk. Cancer Epidemiol Biomarkers Prev. 2008; 17:1203–13. [PubMed: 18483343]

45. Parra EJ, Kittles RA, Argyropoulos G, et al. Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. Am J Phys Anthropol. 2001; 114:18–29. [PubMed: 11150049]

46. Reich DE, Cargill M, Bolk S, et al. Linkage disequilibrium in the human genome. Nature. 2001; 411:199–204. [PubMed: 11346797]

47. Nam RK, Reeves JR, Toi A, et al. A novel serum marker, total prostate secretory protein of 94 amino acids, improves prostate cancer detection and helps identify high grade cancers at diagnosis. J Urol. 2006; 175:1291–7. [PubMed: 16515983]

48. Reeves JR, Dulude H, Panchal C, Daigneault L, Ramnani DM. Prognostic value of prostate secretory protein of 94 amino acids and its binding protein after radical prostatectomy. Clin Cancer Res. 2006; 12:6018–22. [PubMed: 17062675]

49. Yi-Ping F, Wei T, Jacobs K, et al. Refining the prostate cancer genetic association within the JAZF1 gene on chromosome 7p15.2. Cancer Epidemiology, Biomarkers & Prevention. 2010 In Press.

50. Koontz JI, Soreng AL, Nucci M, et al. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. Proc Natl Acad Sci U S A. 2001; 98:6348–53. [PubMed: 11371647]

51. Micci F, Walter CU, Teixeira MR, et al. Cytogenetic and molecular genetic analyses of endometrial stromal sarcoma: nonrandom involvement of chromosome arms 6p and 7p and

confirmation of JAZF1/JJAZ1 gene fusion in t(7;17). Cancer Genet Cytogenet. 2003; 144:119–24. [PubMed: 12850374]

52. Micci F, Panagopoulos I, Bjerkehagen B, Heim S. Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes JAZF1/PHF1 and EPC1/PHF1 in endometrial stromal sarcoma. Cancer Res. 2006; 66:107–12. [PubMed: 16397222]

53. Oliva E, de Leval L, Soslow RA, Herens C. High frequency of JAZF1-JJAZ1 gene fusion in endometrial stromal tumors with smooth muscle differentiation by interphase FISH detection. Am J Surg Pathol. 2007; 31:1277–84. [PubMed: 17667554]

54. Hooker S, Hernandez W, Chen H, et al. Replication of prostate cancer risk loci on 8q24, 11q13, 17q12, 19q33, and Xp11 in African Americans. Prostate. 70:270–5. [PubMed: 19902474]

55. Haiman CA, Le Marchand L, Yamamato J, et al. A common genetic risk factor for colorectal and prostate cancer. Nat Genet. 2007; 39:954–6. [PubMed: 17618282]

56. Robbins C, Torres JB, Hooker S, et al. Confirmation study of prostate cancer risk variants at 8q24 in African Americans identifies a novel risk locus. Genome Res. 2007; 17:1717–22. [PubMed: 17978284]

57. Kader AK, Sun J, Isaacs SD, et al. Individual and cumulative effect of prostate cancer risk-associated variants on clinicopathologic variables in 5,895 prostate cancer patients. Prostate. 2009; 69:1195–205. [PubMed: 19434657]
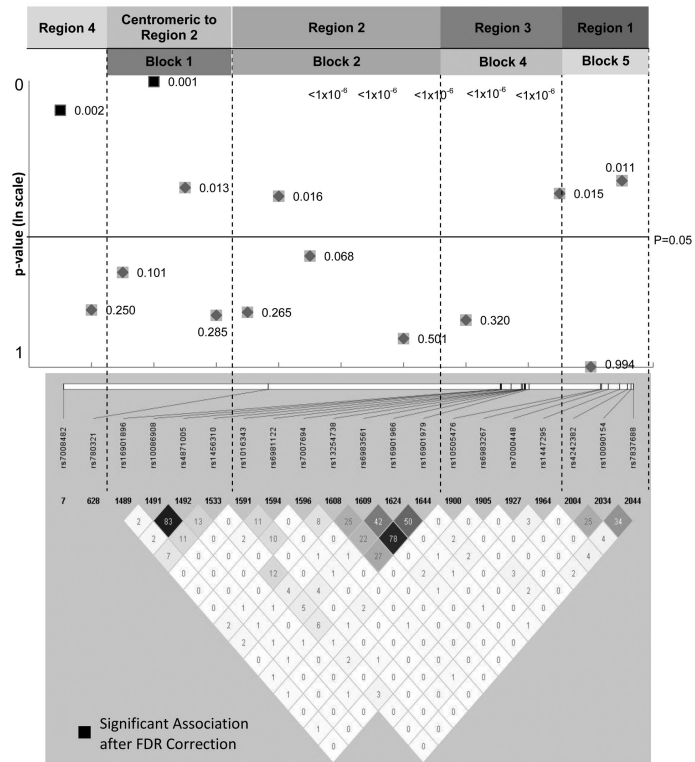
Figure 1.
Results of Prostate Cancer Associations at 8q24 in ADM. P-values for association by genomic location.

Table 1

Results of Associations at Prostate Cancer GWAS Loci in Men of African descent

Chr	Locus	SNP	Position (bp)	Risk Allele	Sample Size		Risk Allele Frequency		aOR (95% CI)*	P-value	OR in EDM***	MAF in EDM**
					Cases	Controls	Cases	Controls				
2	EHBP1	rs721048	62,985,234	A	2083	2116	0.05	0.04	1.02 (0.82 - 1.26)	0.882	1.15(1)	0.19
3	3p12.1	rs2660753	87,193,363	T	1796	1415	0.50	0.49	1.05 (0.95 - 1.17)	0.338	1.18(2)	0.11
6	SLC22A3	rs9364554	160,753,653	T	2652	2659	0.07	0.07	1.05 (0.90 - 1.24)	0.513	1.17(2)	0.29
7	JAZF1	rs10486567	27,943,087	C	2899	2538	0.74	0.71	1.18 (1.08 - 1.29)	0.0002	1.21(3)	0.77
7	LMTK2	rs6465657	97,654,262	C	3266	2631	0.87	0.87	1.03 (0.91 - 1.15)	0.666	1.12(2)	0.46
8	8p21	rs6982080	29,686,860	G	844	536	0.28	0.28	1.08 (0.90 - 1.29)	0.432	1.03(3)	0.31
9	DAB2IP	rs1571801	123,467,193	T	1709	1413	0.15	0.14	1.07 (0.92 - 1.24)	0.385	1.30(4)	0.24
10	MSMB	rs7920517	51,202,626	G	849	541	0.74	0.71	1.13 (0.93 - 1.36)	0.217	1.22(2)	0.48
10	MSMB	rs10993994	51,219,501	T	3374	2982	0.62	0.60	1.12 (1.03 - 1.21)	0.005	1.25(2)	0.4
10	MSMB	rs7904463	51,229,474	T	501	563	0.69	0.66	1.13 (0.92 - 1.39)	0.252	1.06(5)	0.67
10	MINPP1	rs12771728	89,345,292	C	850	541	0.19	0.19	1.03 (0.83 - 1.27)	0.788	1.06(3)	0.34
10	CTBP2	rs4962416	126,686,861	G	1202	1388	0.17	0.17	0.99 (0.85 - 1.16)	0.933	1.18(3)	0.27
11	11q13	rs1241845	68,691,994	A	862	876	0.12	0.13	0.94 (0.76 - 1.16)	0.539	1.16(6)	0.28
11	11q13	rs7931342	68,751,072	G	2445	2018	0.79	0.77	1.15 (1.03 - 1.29)	0.014	1.19(2)	0.49
11	11q13.2	rs10896449	68,751,242	G	2056	1898	0.70	0.68	1.12 (1.01 - 1.24)	0.031	1.17(3)	0.28
12	Chr. 12	rs902774	51,560,170	A	849	541	0.09	0.09	1.02 (0.77 - 1.34)	0.914	1.03(2)	0.15
15	IL16	rs4072111	79,365,193	T	857	539	0.04	0.05	0.87 (0.58 - 1.31)	0.509	1.29(3)	0.11
16	CDH13	rs4782726	81,258,833	A	849	535	0.32	0.32	0.99 (0.83 - 1.19)	0.929	1.20(3)	0.18
17	HNF1B	rs11649743	33,149,091	C	862	874	0.94	0.93	1.10 (0.83 - 1.45)	0.505	1.28(7)	0.8
17	HNF1B	rs4430796	33,172,152	T	3112	2911	0.35	0.33	1.08 (1.00 - 1.18)	0.053	1.22(8)	0.49
17	HNF1B	rs7501939	33,175,268	C	1305	1513	0.50	0.50	0.97 (0.87 - 1.09)	0.651	1.19(8)	0.58
19	KLK2/3	rs887391	46,677,463	T	851	875	0.50	0.49	1.03 (0.90 - 1.18)	0.671	1.13(9)	0.76
19	KLK2/3	rs2735839	56,056,434	A	2773	2680	0.32	0.31	1.07 (0.99 - 1.18)	0.093	1.20(2)	0.15
22	22q13	rs9623117	38,782,064	C	987	1077	0.76	0.76	0.99 (0.86 - 1.15)	0.937	1.25(40)	0.21
X	NUDT10/11	rs5945572	51,246,422	A	1764	1235	0.31	0.28	1.11 (1.02 - 1.20)	0.02	1.23(1)	0.35
X	NUDT10/11	rs5945619	51,258,411	G	1390	1845	0.41	0.36	1.09 (1.00 - 1.18)	0.039	1.19(2)	0.36

*
aOR: per allele OR adjusting for age and study centers.

** OR (per allele OR) in EDM were estimated from: (a) estimations in the original publications whenever they are available (1,2,5,6,7,8,9,19), (b) recalculations based on number of genotype counts in cases and controls when available(4), (c) Specifically for SNPs reported in Thomas et al. (3), ORs for heterozygous carriers were used as a proxy for aORs when OR for homozygous carriers were higher than OR for heterozygous carriers. For SNPs rs4072111 and rs4782726, increased in prostate cancer risk were only seen in homozygous carriers. Therefore, the square-root of OR in homozygous carriers were used as estimates.