Published in final edited form as:

Hum Genet. 2014 March ; 133(3): 347–356. doi:10.1007/s00439-013-1384-2.

Association analysis of 9,560 prostate cancer cases from the International Consortium of Prostate Cancer Genetics confirms the role of reported prostate-cancer associated SNPs for familial disease

Craig C. Teerlink[1], **Stephen N. Thibodeau**[3,4], **Shannon K. McDonnell**[3,5], **Daniel J. Schaid**[3,5], **Antje Rinckleb**[6,7,8], **Christiane Maier**[6,7,8], **Walther Vogel**[6,7], **Geraldine Cancel-Tassin**[9], **Christophe Egrot**[9], **Olivier Cussenot**[9], **William D. Foulkes**[10,11], **Graham G. Giles**[10,12,13], **John L. Hopper**[10,13], **Gianluca Severi**[10,12,13], **Ros Eeles**[10,14], **Douglas Easton**[10,15], **Zsofia Kote-Jarai**[10,14], **Michelle Guy**[10,14], **Kathleen A. Cooney**[16,17], **Anna M. Ray**[16,17], **Kimberly A. Zuhlke**[16,17], **Ethan M. Lange**[16,18], **Liesel M. FitzGerald**[19,20], **Janet L. Stanford**[19,20], **Elaine A. Ostrander**[19,21], **Kathleen E. Wiley**[22], **Sarah D. Isaacs**[22], **Patrick C. Walsh**[22], **William B. Isaacs**[22], **Tiina Wahlfors**[23,24], **Teuvo Tammela**[23,25], **Johanna Schleutker**[23,26], **Fredrik Wiklund**[27,28], **Henrik Grönberg**[27,28], **Monica Emanuelsson**[27,29], **John Carpten**[30], **Joan Bailey-Wilson**[31], **Alice S. Whittemore**[32,33], **Ingrid Oakley-Girvan**[32,34,35], **Chih-Lin Hsieh**[32,36], **William J. Catalona**[37], **S. Lilly Zheng**[37], **Guangfu Jin**[38], **Lingyi Lu**[38], **Jianfeng Xu**[38], **International Consortium for Prostate Cancer Genetics**, **Nicola J. Camp**[1,*], and **Lisa A. Cannon-Albright**[1,2,*]

[1]Division of Genetic Epidemiology, Department of Medicine, University of Utah School of Medicine [2]George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, Utah [3]Mayo Clinic ICPCG Group [4]Department of Lab Medicine and Pathology, Mayo Clinic, Rochester, MN 55905 [5]Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905 [6]University of Ulm ICPCG Group [7]Department of Urology, University of Ulm, Germany [8]Institute for Human Genetics, University of Ulm, Germany [9]CeRePP ICPCG Group, Hopital Tenon, Assistance Publique-Hopitaux de Paris, Paris 75020, France [10]ACTANE Consortium ICPCG Group [11]Program in Cancer Genetics, McGill University, Montreal, Quebec H3T 1E2, Canada [12]Cancer Epidemiology Centre, Cancer Council Victoria, Carlton, VIC 3053, Australia [13]Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, The University of Melbourne, VIC 3010, Australia [14]The Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK [15]Strangeways Laboratory, Worts Causeway, Cambridge CB1 8RN, UK [16]University of Michigan ICPCG Group [17]Departments of Internal Medicine and Urology, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI 48109 [18]Departments of Genetics and Biostatistics, University of North Carolina, Chapel Hill, NC 27599 [19]FHCRC/NHGRI ICPCG Group [20]Fred Hutchinson Cancer Research Center (FHCRC), Division of Public Health Sciences, Seattle, WA 98195 [21]Cancer Genetics Branch, National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH), Bethesda, MD 20892 [22]Johns Hopkins University ICPCG Group, Department of Urology, Johns Hopkins Medical Institutions, Baltimore, MD 21287 [23]University of Tampere ICPCG Group [24]Institute of Biomedical Technology, University of Tampere, BioMediTech, and Centre for Laboratory Medicine, Tampere University Hospital, Tampere 33520, Finland [25]Department of Urology, University of Tampere and Tampere University Hospital, Tampere 33520, Finland [26]Department of Medical Biochemistry and Genetics, University of Turku, Turku 20520, Finland [27]University of Umeå ICPCG Group

Corresponding Author: Craig C. Teerlink, PhD, craig.teerlink@utah.edu, Tel: 801.587.9303, Fax: 801.581.6052.
*Equal contribution.

[28]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden [29]Oncologic Centre, Umeå University, Umeå 90187, Sweden [30]Integrated Cancer Genomics Division, TGen, Phoenix, AZ 85004 [31]Inherited Disease Research Branch, National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH), Bethesda, MD 20892 [32]Stanford University ICPCG Group [33]Department of Health Research and Policy, Stanford School of Medicine, Stanford, CA 94305 [34]Cancer Prevention Institute of California, 2201 Walnut Ave Suite 300, Fremont, CA 94538 [35]Stanford Cancer Institute and Department of Health Research and Policy, Stanford School of Medicine, Stanford, CA 94305 [36]Department of Urology and Department of Biochemistry and Molecular Biology, University of Southern California, Los Angeles, CA 90089 [37]Northwestern University ICPCG Group, Northwestern University Feinberg School of Medicine, Chicago, IL 60611 [38]Data Coordinating Center for the ICPCG and Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157

Abstract

Previous GWAS studies have reported significant associations between various common SNPs and prostate cancer risk using cases unselected for family history. How these variants influence risk in familial prostate cancer is not well studied. Here, we analyzed 25 previously reported SNPs across 14 loci from prior prostate cancer GWAS. The International Consortium for Prostate Cancer Genetics (ICPCG) previously validated some of these using a family-based association method (FBAT). However, this approach suffered reduced power due to the conditional statistics implemented in FBAT. Here, we use a case-control design with an empirical analysis strategy to analyze the ICPCG resource for association between these 25 SNPs and familial prostate cancer risk. Fourteen sites contributed 12,506 samples (9,560 prostate cancer cases, 3,368 with aggressive disease, and 2,946 controls from 2,283 pedigrees). We performed association analysis with Genie software which accounts for relationships. We analyzed all familial prostate cancer cases and the subset of aggressive cases. For the familial prostate cancer phenotype, 20 of the 25 SNPs were at least nominally associated with prostate cancer and 16 remained significant after multiple testing correction ($p \leq 1E^{-3}$) occurring on chromosomal bands 6q25, 7p15, 8q24, 10q11, 11q13, 17q12, 17q24, and Xp11. For aggressive disease, 16 of the SNPs had at least nominal evidence and 8 were statistically significant including 2p15. The results indicate that the majority of common, low-risk alleles identified in GWAS studies for all prostate cancer also contribute risk for familial prostate cancer, and that some may be contribute risk to aggressive disease.

Keywords

prostate cancer; pedigrees; familial disease; simulation; replication

Introduction

Previous prostate cancer GWAS have reported associations between various SNPs and prostate cancer in cohorts of prostate cancer cases unselected for family history (Amundadottir 2006; Duggan 2007; Gudmundsson 2007a; Gudmundsson 2007b; Haiman 2007; Eeles 2008; Gudmundsson 2008; Salinas 2008; Sun 2008; Thomas 2008). The International Consortium for Prostate Cancer Genetics (ICPCG) selected 25 of these SNPs to pursue replication of these findings in a set of related hereditary prostate cancer cases selected for membership in high-risk pedigrees. A previous analysis of the ICPCG data used family based association testing (FBAT) on 102 – 477 informative families and was able to confirm three of these candidate SNPs ($p \leq 2E^{-3}$ (= 0.05/25)) (Jin 2012). Here a larger analysis of the same 25 SNPs in over 12,000 individuals was conducted using a case-control framework that allowed analysis of all data submitted by ICPCG member sites without

restriction to the trio relationship structure. The increased sample size considerably improves statistical power to study these SNPs.

Fourteen study sites contributed a total of 12,506 samples for genotyping, including 2,946 controls, 6,192 cases with non-aggressive disease, and 3,368 cases with aggressive disease. Genotyped samples originated from 2,283 pedigrees. Each site contributed its own controls, with an average of 231 controls per site, except for one site that provided genotype data for 931 genetically matched publicly available controls. It is well known that close relationships can have an inflationary effect on statistics for tests of association, therefore, it was necessary to account for known relationships in the analysis. Genie software was used to accomplish this (Allen-Brady 2005; Curtin 2007). Genie generates an empirical null distribution, matched for the known pedigree structures and multiple sites, from which to assess the observed test statistic for significance. In this study 10 million such simulations were used to estimate the necessary null distributions. Separate analyses for all familial prostate cancers and the subset of aggressive prostate cancers were conducted.

Methods and Materials

Sample cohort

Fourteen member sites of the ICPCG consortium provided samples for analysis; these sites were the African American Hereditary Prostate Cancer Consortium (AAHPC), the Anglo/Canadian/Texan/Australian/Norwegian/European Union Biomed (ACTANE), University of Tampere (Finland), Fred Hutchinson Cancer Research Center (FHCRC), Centre de Recherche pour les Pathologies Prostatiques (France), Johns Hopkins University (JHU), the Mayo Clinic (Mayo), The University of Michigan (Michigan), The University of Montreal (Montreal), Northwestern University (NW), Stanford University (Stanford), University of Umea (Sweden), University of Ulm (Ulm), and University of Utah (Utah). Each site recruited study participants according to their own protocols; however, for consistency, confirmation from either death certificate or medical records was required for a diagnosis of prostate cancer.

Table 1 provides the number of cases analyzed from each site. Each site also provided control samples, which were: unaffected pedigree members; regionally selected and ethnically matched controls; or (for one site) in silico controls. The in silico controls were supplied by Michigan, who provided 931 controls from the Illumina Genotype Control Database (iControlDB) (www.illumina.com) that had been genetically matched to their set of cases using 610K SNPs (Genomic Inflation Factor (Clayton 2005) of 1.018). Several sites also provided population ascertained cases (non-familial) that were used as another comparator group in a secondary genetic risk score analysis (n = 1,872).

All cases analyzed in this study from all study sites were of Caucasian ethnic background with the exception of the cases supplied by the AAHPC site which were all from African American pedigrees.

Phenotypes

In order to address the potential for clinical heterogeneity across all sites, we imposed standardized criteria for prostate cancer status. All prostate cancer cases were confirmed by death or medical record. Each was designated as non-aggressive, aggressive, or undetermined aggressive status. Cases were considered aggressive if they were categorized as regional or distant stage, poorly differentiated or non-differentiated grade, or had evidence for death due to metastatic prostate cancer (Schaid 2006; Christensen 2007). Two phenotypes were analyzed. The first phenotype consisted of all prostate cancer cases in the pedigrees regardless of aggressiveness. A separate analysis of aggressive cases only was

also performed; with comparisons made both to controls and to all non-aggressive prostate cases.

Genotypes

SNPs were selected for genotyping based on previously published reports that an allele at the SNP was significantly associated with prostate cancer; SNPs are shown in Table 2. These SNPs occurred at cytogenetic bands 2p15, 3p12, 6q25, 7p15, 7q21, 8q24, 9q33, 10q11, 10q26, 11q13, 17q12, 17q24, 19q13 and Xp11. Genotyping was performed with MassARRAY iPLEX (Sequenom, Inc., San Diego, CA) at the Center for Cancer Genomics, Wake Forest University and is further described elsewhere (Jin 2012). Since imputation of missing genotypes is not possible given the paucity of genotyped SNPs, individuals with missing data at particular SNPs were ignored in those analyses.

Statistical Methods

The vast majority of the cases analyzed in this study reside in high-risk pedigrees. It is well known that standard association techniques are not appropriate for related individuals due to lack of independence of genotypes. Analyses were conducted with Genie software, which allows for valid analysis of all data, whether independent or not. To account for relatedness, Genie software compares the observed test statistic to an empirical null distribution derived from simulated data sets matched for pedigree structure but generated under the null hypothesis. In brief, the pedigree founders are assigned alleles based on their population frequencies and alleles of subsequent pedigree members are assigned according to random Mendelian inheritance of the founder alleles (a ‘gene-drop’). Test statistics are calculated for each null simulation to determine a null distribution from which the observed statistic can be assessed and an empirical p-value is assigned. Singleton cases/controls are simply considered as founders with no descendants. For multi-site analyses, simulations are generated in a site-specific manner and overall association evidence is based on a Cochran-Mantel-Haenzsel meta-statistic across sites (Mantel 1959; Agresti 1990; Curtin 2007). The primary analysis was the allele test for association for each of the 25 SNPs (asymptotically equivalent to a trend test). Up to 10 million simulations were used in the null distribution to estimate p-values, depending on the necessary resolution required. Analysis began with 10,000 simulations and for SNPs with an empirical p-value ≤ 0.1, an additional 10x number of simulations were performed; this process was repeated until a maximum of 10 million simulations were performed. A significance threshold of $p \leq 1E^{-3}$ was used to declare statistical significance accounting for multiple tests, which represents a Bonferroni corrected p-value for 25 tests and 2 phenotypes (corrected alpha = 0.05/50). A Q-test was used to identify SNPs that exhibited significant heterogeneity across sites.

As follow-up, secondary to the main effects analyses described above, three additional analyses were performed using Genie. First, we tested all two-way interactions between all pairs of SNPs (assessed by significance of the interaction coefficient in a logistic regression framework). Second, we compared aggressive cases to non-aggressive cases by recoding aggressive cases as ‘cases’ and non-aggressive cases as ‘controls’ and performed tests of association at each marker. Third, we estimated a genetic risk score based on the number of risk alleles carried across the SNPs identified as significant in the main analyses. To avoid sparse data for the number of risk alleles, the extremes of the scale were collapsed to contain the top/bottom 5% of the data, the resulting categories were: 0–8, 9, 10, …, 16, 17, 18–32 risk alleles carried. Only individuals with genotype data at all markers were included in the analysis. A trend test across these groups, weighted by the number of risk alleles, was used to compare the distributions for cases and controls. The genetic risk score test was repeated using the non-familial cases supplied by several of the contributing sites in order to establish

the extent to which the genetic risk may differ between familial and non-familial prostate cancer cases.

Results

Table 3 shows the results of the primary analyses: meta-analysis combining data from all sites for the phenotypes consisting of all prostate cancer cases and for aggressive cases only. For the prostate cancer phenotype, 20 of the 25 SNPs were nominally significant ($p \leq 0.05$), and 16 remained statistically significant after correction for multiple testing (p-values ranging from $1E^{-3}$ to $\leq 1E10^{-7}$). Replicated SNPs were on chromosomal bands 6q25, 7p15, 8q24, 10q11, 11q13, 17q12, 17q24, and Xp11. The odds ratios for all but 1 of the significant SNPs were less extreme than the originally published findings, as is often the situation in replication studies, although perhaps surprising for familial cases (Table 3).

For the aggressive prostate cancer phenotype, 16 markers showed at least nominal evidence and 8 of the 25 SNPs were statistically significant after correcting for multiple testing (p-values ranging from $1.3E^{-3}$ to $\leq 1E10^{-7}$). Qualitatively, the results for the aggressive phenotype were similar to the results of all prostate cancer cases indicating that these SNPs do not offer substantial discrimination between these two clinically distinguishable phenotypes. For any SNPs that were significant for both phenotypes, the odds ratio was consistently more extreme for aggressive prostate cancer; although none were significantly different. In accordance with this, in the secondary analysis of aggressive versus non-aggressive disease, no statistically significant differences were found (results not shown).

The results for the 14 study sites are reported in Supplemental Figure 1 for all prostate cancer and in Supplemental Figure 2 for aggressive prostate cancer, depicted in Forrest plots. Four SNPs indicated significant heterogeneity across sites (Q-test $p \leq 1E^{-3}$); two at 8q24 (rs1447295; rs10090154), 9q33 (rs1571801), and Xp11 (rs5945619). Three of these 4 SNPs were significantly replicated, and by inspection of the Forrest plots it can be seen that the by-site odds ratio estimates vary with one or two sites having extreme risk estimates but in the same direction as the meta-analysis result. The fourth SNP, at rs1571801 on 9q33, was not significant in the meta-analysis. It is notable that for this SNP the AAHPC site had a by-site significant OR estimate (OR = 1.7; 95% CI = (1.2, 2.5)) which was in the opposite direction from most of the other sites. The odds ratio estimate for the aggressive phenotype for this site at this marker was even more extreme (OR = 2.4; 95% CI = (1.5, 4.4)). The AAHPC site differs from the other sites in that it is composed of African American families, indicating that this SNP may have a role in familial prostate cancer for this ethnic group even though the marker failed to achieve significance overall. The SNP rs1571801 is an intronic polymorphism in the DAB2IP gene, a documented tumor suppressor that has been observed to be aberrantly methylated in some prostate and lung cancers (Yano 2005) and has been associated with early onset prostate cancer in a set of 754 Caucasians (Lange 2012). According to 1000 Genomes project, the minor allele frequency of this SNP does exhibit some variation between ethnic groups (5% in Asians; 15% in Africans; 20% in ad-mixed Americans; 24% in Europeans), indicating that this SNP may be in linkage disequilibrium with some causal variant(s) in the African American prostate cancer families in this study, but not in the familial cases studied with other ancestral backgrounds.

The analysis of all two-way interactions failed to identify any statistically significant interactions, after adjusting for multiple testing.

The results of the genetic risk analysis appear in Table 4. The Table shows the ORs and 95% CIs of familial cases and controls comparing the number of risk alleles carried across the 16 replicated SNPs. In a test for trend weighted by the number of risk alleles, familial prostate

cancer cases were significantly different than controls ($p \leq 1E^{-6}$). The results presented in Table 4 show that several adjacent categories of risk alleles carried exhibit similar levels of risk and could be collapsed. For instance, using a baseline of ≤8 risk alleles (lowest 10% of distribution), the increased risk for familial cases was approximately 1.6 for 9 or 10 risk alleles (19% of the population), 1.93 for 11–13 risk alleles (40%), and 2.84 for 14+ alleles (31%). It is notable that when taken together, these 16 risk loci appear to be able to distinguish extreme groups that could be clinical valuable for determining early or more frequent screening for prostate cancer prevention. This analysis was repeated for population cases (results appear in Supplemental Table 1) and also found to be highly significant ($p \leq 1E^{-6}$). The increased risks were estimated to be approximately 1.40 for 9 or 10 risk alleles, 1.70 for 11–13 risk alleles, and 2.74 for 14+ alleles. The difference between familial and population cases was statistically significant ($p \leq 1E^{-6}$), with familial cases carrying more risk alleles, further confirming that family history is an important determinant in prostate risk.

Discussion

This familial case-control analysis of 9,560 familial prostate cancer cases significantly confirmed 16 of 25 SNPs ($p \leq 1E^{-3}$) previously reported to be associated with prostate cancer in population-based GWAS. This is compared to only three SNPs that could be replicated at the same significance using a FBAT analysis nested in the same genotype data (Jin 2012). The clear advantage of this analysis strategy (using related familial cases in a traditional case-control design) is the ability to use all the available genotype data. This produced a notable enhancement to statistical power. With the increased sample size, there was sufficient power for a subset analysis for aggressive disease, for which association of a SNP in the 2p15 locus for these familial cases was validated. Specifically, beyond the three loci with significant evidence from the FBAT analysis (10q11, 17q24 and Xp11), significant replication evidence for SNPs at 5 additional loci was shown: 2q15 (aggressive disease only), 7p15, 8q24, 11q13, and 17q12.

The results of this analysis highlight the benefit of using the case-control design, even with family-based data. The approach reported here was enabled by the flexible empirical approach contained in the Genie software that can appropriately account for the relatedness among cases. Furthermore, Genie software also provides a valid means to adjust for site-specific effects in a meta-analysis framework, and the ability to test for interaction effects and multi-locus analyses (such as a genetic risk score).

An analysis of familial aggressive cases versus familial non-aggressive prostate cancer cases was performed, but did not identify any statistically significant differences. This outcome indicates that, at least for the definition of aggressiveness that we used, these 25 SNPs do not distinguish between risk for aggressive and non-aggressive prostate cancer. The fact that two regions failed to achieve significance in the aggressive phenotype analysis but did achieve significance in the all-PRCA phenotype analysis (6q25 and 7p15) indicates that this outcome is more likely an artifact of diminishing sample size and not due to a clinically important mechanism between the two disease definitions (the two analyses provided very similar odds ratios).

The genetic risk score based on the 16 replicated SNPs was an attempt to consider the multi-locus combined risk across multiple disease-associated SNPs. This analysis revealed that, although the individual SNP ORs were less extreme than the initial reports, that considering all 16 loci together, the familial prostate cancer cases carry significantly more risk alleles than do sporadic cases, indicating that they are enriched for these genetic associations. This outcome indicates that these families may be of great value for sequencing efforts to identify

the genetic factors underlying these associations. Indeed, the ICPCG has already made a strong effort towards that objective. This observation of stronger genetic risk in familial disease is consistent with a recent report that positive family history doubles lifetime risk of prostate cancer above that attributable to carrying all risk alleles across 26 common variants (MacInnis 2011). Another recent report of a genetic risk score analysis comparing population ascertained cases and controls compared genotypes collected on 33 common variants previously shown to be associated with prostate cancer and demonstrated only a marginal improvement in prostate cancer prediction over prostate specific antigen screening alone (Johansson 2012). This question remains to be answered for familial prostate cancer where the effects appear to be larger.

In conclusion, these observations support that the majority of SNPs identified from GWAS using population-based case-control cohorts likely also play a role in the risk of familial disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Partial support for L.A.C.A. and for all data sets with in the Utah Population Database (UPDB) was provided by Huntsman Cancer Institute, University of Utah and the Huntsman Cancer Institute's Cancer Center Support grant, P30 CA42014 from National Cancer Institute. Research was supported by the Utah Cancer Registry, which is funded by Contract No. **HHSN261201000026C** from the National Cancer Institute's SEER Program with additional support from the Utah State Department of Health and the University of Utah. A subcontract from Johns Hopkins University with funds provided by grant R01 CA89600 from the NIH National Cancer Institute (to L.A. Cannon Albright). The International Consortium for Prostate Cancer Genetics has provided access to genotyping for this study (U01CA089600). RE is supported by Cancer Research UK and Prostate Action (now Prostate Cancer UK) and National Institute of Health Research support to the Biomedical Research Centre at The Institute of Cancer Research and Royal Marsden NHS Foundation Trust. Robert Stephenson generously donated funds for the data analysis aspects of this project.

References

1. Agresti, A. Categorical Data Analysis. John Wiley & Sons, Inc; New York: 2001.
2. Allen-Brady K, Wong J, Camp NJ. PedGenie: an analysis approach for genetic association testing in extended pedigrees and genealogies of arbitrary size. BMC Cancer. 2005; 5:99. [PubMed: 16091150]
3. 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, Bälter 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; 38:652–658. [PubMed: 16682969]
4. Christensen GB, Camp NJ, Farnham JM, Cannon-Albright LA. Genome-wide linkage analysis for aggressive prostate cancer in Utah high-risk pedigrees. Prostate. 2007; 67:605–613. [PubMed: 17299800]
5. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, Smink LJ, Lam AC, Ovington NR, Stevens HE, Nutland S, Howson JM, Faham M, Moorhead M, Jones HB, Falkowski M, Hardenbol P, Willis TD, Todd JA. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet. 2005; 37:1243–1246. [PubMed: 16228001]
6. Curtin K, Wong J, Allen-Brady K, Camp NJ. PedGenie: Meta genetic association testing in mixed family and case-control designs. BMC Bioinformatics. 2007; 15:448. [PubMed: 18005446]

7. Duggan D, Zheng SL, Knowlton M, Benitez D, Dimitrov L, Wiklund F, Robbins C, Isaacs SD, Cheng Y, Li G, Sun J, Chang BL, Marovich L, Wiley KE, Bälter K, Stattin P, Adami HO, Gielzak M, Yan G, Sauvageot J, Liu W, Kim JW, Bleecker ER, Meyers DA, Trock BJ, Partin AW, Walsh PC, Isaacs WB, Grönberg H, Xu J, Carpten JD. Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. J Natl Cancer Inst. 2007; 99:1836–1844. [PubMed: 18073375]
8. 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. UK Genetic Prostate Cancer Study Collaborators; British Association of Urological Surgeons' Section of Oncology; UK ProtecT Study Collaborators. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet. 2008; 40:316–321. [PubMed: 18264097]
9. Gudmundsson J, Sulem P, Rafnar T, Bergthorsson JT, Manolescu A, Gudbjartsson D, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Blondal T, Jakobsdottir M, Stacey SN, Kostic J, Kristinsson KT, Birgisdottir B, Ghosh S, Magnusdottir DN, Thorlacius S, Thorleifsson G, Zheng SL, Sun J, Chang BL, Elmore JB, Breyer JP, McReynolds KM, Bradley KM, Yaspan BL, Wiklund F, Stattin P, Lindström S, Adami HO, McDonnell SK, Schaid DJ, Cunningham JM, Wang L, Cerhan JR, St Sauver JL, Isaacs SD, Wiley KE, Partin AW, Walsh PC, Polo S, Ruiz-Echarri M, Navarrete S, Fuertes F, Saez B, Godino J, Weijerman PC, Swinkels DW, Aben KK, Witjes JA, Suarez BK, Helfand BT, Frigge ML, Kristjansson K, Ober C, Jonsson E, Einarsson GV, Xu J, Gronberg H, Smith JR, Thibodeau SN, Isaacs WB, Catalona WJ, Mayordomo JI, Kiemeney LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet. 2008; 40:281–283. [PubMed: 18264098]
10. 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. 2007a; 39:631–637. [PubMed: 17401366]
11. 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. 2007b; 39:977–983. [PubMed: 17603485]
12. 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; 39:638–644. [PubMed: 17401364]

13. Jin G, Lu L, Cooney KA, Ray AM, Zuhlke KA, Lange EM, Cannon-Albright LA, Camp NJ, Teerlink CC, Fitzgerald LM, Stanford JL, Wiley KE, Isaacs SD, Walsh PC, Foulkes WD, Giles GG, Hopper JL, Severi G, Eeles R, Easton D, Kote-Jarai Z, Guy M, Rinckleb A, Maier C, Vogel W, Cancel-Tassin G, Egrot C, Cussenot O, Thibodeau SN, McDonnell SK, Schaid DJ, Wiklund F, Grönberg H, Emanuelsson M, Whittemore AS, Oakley-Girvan I, Hsieh CL, Wahlfors T, Tammela T, Schleutker J, Catalona WJ, Zheng SL, Ostrander EA, Isaacs WB, Xu J. International Consortium for Prostate Cancer Genetics. Validation of prostate cancer risk-related loci identified from genome-wide association studies using family-based association analysis: evidence from the International Consortium for Prostate Cancer Genetics (ICPCG). Hum Genet. 2012; 131:1095–1103. [PubMed: 22198737]

14. Johansson M, Holmström B, Hinchliffe SR, Bergh A, Stenman UH, Hallmans G, Wiklund F, Stattin P. Combining 33 genetic variants with prostate-specific antigen for prediction of prostate cancer: longitudinal study. Int J Cancer. 2012; 130:129–137. [PubMed: 21328341]

15. Macinnis RJ, Antoniou AC, Eeles RA, Severi G, Al Olama AA, McGuffog L, Kote-Jarai Z, Guy M, O'Brien LT, Hall AL, Wilkinson RA, Sawyer E, Ardern-Jones AT, Dearnaley DP, Horwich A, Khoo VS, Parker CC, Huddart RA, Van As N, McCredie MR, English DR, Giles GG, Hopper JL, Easton DF. A risk prediction algorithm based on family history and common genetic variants: application to prostate cancer with potential clinical impact. Genet Epidemiol. 2011; 35:549–556. [PubMed: 21769933]

16. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959; 22:719–748. [PubMed: 13655060]

17. 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; 17:1203–1213. [PubMed: 18483343]

18. Schaid DJ, McDonnell SK, Zarfas KE, Cunningham JM, Hebbring S, Thibodeau SN, Eeles RA, Easton DF, Foulkes WD, Simard J, Giles GG, Hopper JL, Mahle L, Moller P, Badzioch M, Bishop DT, Evans C, Edwards S, Meitz J, Bullock S, Hope Q, Guy M, Hsieh CL, Halpern J, Balise RR, Oakley-Girvan I, Whittemore AS, Xu J, Dimitrov L, Chang BL, Adams TS, Turner AR, Meyers DA, Friedrichsen DM, Deutsch K, Kolb S, Janer M, Hood L, Ostrander EA, Stanford JL, Ewing CM, Gielzak M, Isaacs SD, Walsh PC, Wiley KE, Isaacs WB, Lange EM, Ho LA, Beebe-Dimmer JL, Wood DP, Cooney KA, Seminara D, Ikonen T, Baffoe-Bonnie A, Fredriksson H, Matikainen MP, Tammela TL, Bailey-Wilson J, Schleutker J, Maier C, Herkommer K, Hoegel JJ, Vogel W, Paiss T, Wiklund F, Emanuelsson M, Stenman E, Jonsson BA, Grönberg H, Camp NJ, Farnham J, Cannon-Albright LA, Catalona WJ, Suarez BK, Roehl KA. Pooled genome linkage scan of aggressive prostate cancer: results from the International Consortium for Prostate Cancer Genetics. Hum Genet. 2006; 120:471–485. [PubMed: 16932970]

19. Sun J, Zheng SL, Wiklund F, Isaacs SD, Purcell LD, Gao Z, Hsu FC, Kim ST, Liu W, Zhu Y, Stattin P, Adami HO, Wiley KE, Dimitrov L, Sun J, Li T, Turner AR, Adams TS, Adolfsson J, Johansson JE, Lowey J, Trock BJ, Partin AW, Walsh PC, Trent JM, Duggan D, Carpten J, Chang BL, Grönberg H, Isaacs WB, Xu J. Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. Nat Genet. 2008; 40:1153–1155. [PubMed: 18758462]

20. 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; 40:310–315. [PubMed: 18264096]

21. Yano M, Toyooka S, Tsukuda K, Dote H, Ouchida M, Hanabata T, Aoe M, Date H, Gazdar AF, Shimizu N. Aberrant promoter methylation of human DAB2 interactive protein (hDAB2IP) gene in lung cancers. Int J Cancer. 2005; 113(1):59–66. [PubMed: 15386433]

Table 1

Composition of study cohort by contributing sites. Counts reflect the maximum number of individuals genotyped at one marker for each site.

	Familial ascertainment			Population-based ascertainment	
Site	**No. of pedigrees**	**PRCA cases (all)**	**aggressive PRCA cases**	**PRCA cases**	**Unaffected**[a]
AAHPC	81	213	37	*0*	240
ACTANE	304	779	167	*0*	104
FHCRC	270	1061	349	*171*	371
FINLAND	76	421	228	*178*	485
FRANCE	160	666	302	*154*	225
JHU	220	899	248	*0*	231
MAYO	186	1138	664	*640*	563
MICHIGAN	317	1186	402	*0*	1109[b]
MONTREAL	0	0	0	*167*	0
NW	28	128	42	*26*	40
STANFORD	98	253	43	*0*	62
SWEDEN	91	292	120	*0*	91
ULM	190	849	430	*319*	300
UTAH	262	1675	336	*217*	234
Total:	2283	9560	3368	1872	2946

[a] Unaffecteds originated from either familial and population-based ascertainment

[b] 930 controls for this site originated from the illumina icontrols database

Table 2

Description of analyzed markers.

Cytogenetic band	SNP	bp position (hg19)	Gene	Non-risk/risk	Risk allele frequency
2p15	rs721048	63,131,481	EHBP1	G/A	0.178
3p12	rs2660753	87,110,424		C/T	0.11
6q25	rs9364554	160,833,414	SLC22A3	C/T	0.309
7p15	rs10486567	27,976,313	JAZF1	T/C	0.778
7q21	rs6465657	97,816,077	LMTK2	T/C	0.491
8q24	rs979200	127,923,470		T/C	0.66
8q24	rs1016343	128,093,047		C/T	0.245
8q24	rs13254738	128,104,093		A/C	0.341
8q24	rs6983561	128,106,630		A/C	0.06
8q24	rs16901979	128,124,666		C/A	0.062
8q24	rs6983267	128,413,555		T/G	0.561
8q24	rs7837328	128,422,877		G/A	0.452
8q24	rs7000448	128,440,920		C/T	0.396
8q24	rs1447295	128,484,788		C/A	0.163
8q24	rs4242382	128,517,323		G/A	0.156
8q24	rs10090154	128,531,887		C/T	0.155
9q33	rs1571801	124,427,123	DAB2IP	G/T	0.278
10q11	rs10993994	51,549,246	MSMB	C/T	0.427
10q26	rs4962416	126,696,622	CTBP2	A/G	0.267
11q13	rs10896449	68,994,417		A/G	0.549
17q12	rs11649743	36,074,729	HNF1B	T/C	0.838
17q12	rs4430796	36,097,790	HNF1B	C/T	0.59
17q24	rs1859962	69,108,503		T/G	0.526
19q13	rs2735839	51,364,373	KLK3	A/G	0.869
Xp11	rs5945619	51,241,422		A/G	0.428

Table 3

Allele test results for meta-analysis of 25 SNPs for all prostate cancer cases and the subset of aggressive prostate cancer cases. Empirical p-values were estimated with up to 10 million simulations. Bold text denotes statistical significance ($p \leq 1E^{-3}$).

SNP	All PrCa Meta Q-test p-value	All PrCa Meta analysis (OR, EMP P, 95% CI)	Aggressive PrCa Meta Q-test p-value	Aggressive PrCa Meta analysis (OR, EMP P, 95% CI)	Originally reported Odds ratios (OR,95% CI)
rs721048	5.37E-01	1.15, 3E-3, (1.05, 1.26)	5.87E-01	1.25, 1.06E-3, (1.09, 1.42)	1.15 (1.10, 1.21)[a]
rs2660753	7.63E-01	1.08, 1.1E-1, (0.98, 1.19)	8.39E-01	1.11, 2.1E-1, (0.94, 1.3)	1.52 (1.30, 1.77)[b]
rs9364554	6.73E-01	**1.15, 1.34E-4, (1.07, 1.24)**	6.00E-01	1.09, 1.38E-1, (0.97, 1.22)	1.28 (1.16, 1.41)[b]
rs10486567	7.01E-01	**1.2, 4E-6, (1.11, 1.3)**	1.40E-01	1.22, 2.11E-3, (1.08, 1.39)	1.35 (1.20, 1.51)[c,*]
rs6465657	9.70E-02	1.08, 4E-2, (1, 1.16)	4.36E-01	1.05, 3.99E-1, (0.94, 1.16)	1.30 (1.19, 1.43)[b]
rs979200	2.80E-01	1.06, 8E-2, (0.99, 1.14)	7.75E-01	1.09, 1.1E-1, (0.98, 1.22)	0.76 (0.64, 0.90)[d,*]
rs1016343	7.71E-01	**1.19, 4E-6, (1.11, 1.29)**	7.36E-01	1.15, 2.43E-2, (1.02, 1.3)	1.32 (1.12, 1.57)[d,*]
rs13254738	3.72E-01	1.01, 7.32E-1, (0.94, 1.1)	2.86E-01	1.02, 7.1E-1, (0.91, 1.14)	1.26 (1.18, 1.36)[e]
rs6983561	7.25E-03	**1.36, 1.28E-5, (1.19, 1.56)**	7.75E-03	1.29, 2.2E-2, (1.04, 1.6)	1.42 (1.28, 1.58)[e]
rs16901979	1.65E-02	**1.31, 7.86E-6, (1.11, 1.52)**	3.64E-02	1.39, 2.72E-3, (1.12, 1.73)	1.79 (1.53, 2.11)[f]
rs6983267	2.46E-01	**1.20, <1E-7, (1.12, 1.28)**	3.87E-01	1.13, 2.02E-2, (1.02, 1.26)	1.18 (1.09, 1.27)[e]
rs7837328	2.13E-01	**1.14, 6.4E-5, (1.07, 1.22)**	9.90E-02	1.11, 5.37E-2, (1, 1.23)	1.26 (1.06, 1.50)[d]
rs7000448	8.69E-01	1.09, 2.6E-2, (1.01, 1.17)	9.18E-01	1.13, 2.77E-2, (1.01, 1.26)	1.26 (1.15, 1.38)[e]
rs1447295	3.00E-04	**1.39, <1E-7, (1.27, 1.52)**	1.57E-01	**1.46, <1E-7, (1.27, 1.69)**	1.72 (1.33, 2.18)[g]
rs4242382	2.20E-03	**1.39, <1E-7, (1.27, 1.52)**	4.10E-02	**1.39, 4E-6, (1.21, 1.61)**	1.41 (1.24, 1.60)[c,*]
rs10090154	1.00E-03	**1.36, <1E-7, (1.23, 1.5)**	8.90E-02	**1.47, <1E-7, (1.27, 1.7)**	1.43 (1.30, 1.58)[e]
rs1571801	1.00E-03	1.01, 7.92E-1, (0.94, 1.08)	1.80E-03	0.99, 9.04E-1, (0.88, 1.12)	1.27 (1.10, 1.48)[h]
rs10993994	1.26E-01	**1.20, <1E-7, (1.12, 1.28)**	1.96E-01	**1.20 9.07E-4, (1.08, 1.33)**	1.24 (1.10, 1.39)[c,*]
rs4962416	5.92E-01	1.08, 3.6E-2, (1, 1.16)	6.68E-01	1.05, 4.25E-1, (0.93, 1.18)	1.20 (1.07, 1.34)[c,*]
rs10896449	1.84E-01	**1.17, 4.8E-6, (1.09, 1.25)**	2.54E-01	**1.28, 7.2E-6, (1.15, 1.43)**	1.28 (1.13, 1.45)[c,*]
rs11649743	7.40E-02	**1.17, 8.8E-4, (1.06, 1.28)**	1.60E-02	**1.34, 4.08E-5, (1.17, 1.54)**	1.28 (1.07, 1.52)[i,*]
rs4430796	2.70E-02	**1.12, 5.2E-4, (1.05, 1.2)**	1.19E-01	1.17, 3.09E-3, (1.06, 1.31)	1.22 (1.15, 1.30)[j]
rs1859962	4.92E-01	**1.17, 2.8E-6, (1.09, 1.25)**	1.90E-01	**1.24, 5.48E-5, (1.12, 1.38)**	1.20 (1.14, 1.27)[j]

SNP	All PrCa Meta Q-test p-value	All PrCa Meta analysis (OR, EMP P, 95% CI)	Aggressive PrCa Meta Q-test p-value	Aggressive PrCa Meta analysis (OR, EMP P, 95% CI)	Originally reported Odds ratios (OR,95% CI)
rs2735839	4.82E-01	1.16, 1.36E-3, (1.06, 1.27)	4.50E-01	1.07, 4.01E-1, (0.92, 1.24)	0.56 (0.50, 0.64)[b]
rs5945619	< 1E-04	**1.23, 6.8E-5, (1.11, 1.36)**	< 1E-04	**1.28, 4.8E-5, (1.13, 1.44)**	1.46 (1.28, 1.66)[b]

Abbreviations. OR: odds ratio; EMP P: empirical p-value; PrCa: prostate cancer.

[*] Heterogeneity odds ratio was reported.

[a] (Gudmundsson 2008);

[b] (Eeles 2008);

[c] (Thomas 2008);

[d] (Salinas 2008);

[e] (Haiman 2007);

[f] (Gudmundsson 2007a);

[g] (Amundadottir 2006);

[h] (Duggan 2007);

[i] (Sun 2008);

[j] (Gudmundsson 2007b).

Table 4

Genetic risk score analysis for familial cases versus all available controls. Odds ratios (ORs), empirically estimated p-values (from 100,000 simulations) and 95% confidence intervals are given. Counts represent the total number of risk alleles carried. The odds ratios are calculated for each quantity of risk alleles carried compared to the first category, which includes 0–8 risk alleles. Empirical p-values estimated from 100,000 simulations.

Risk alleles carried	No. of cases	Percent of cases	No. of controls	Percent of controls	OR	EMP P	95% CI
0–8	394	7.67%	352	14.72%	-	-	-
9	374	7.28%	231	9.66%	1.5	9.7E-03	(1.1, 2.0)
10	527	10.26%	279	11.66%	1.6	8.4E-04	(1.2, 2.1)
11	690	13.43%	330	13.80%	1.8	2.0E-05	(1.4, 2.3)
12	740	14.40%	289	12.08%	2.2	< 1E-5	(1.7, 2.9)
13	660	12.85%	314	13.13%	1.8	< 1E-5	(1.4, 2.4)
14	557	10.84%	192	8.03%	2.6	< 1E-5	(2.0, 3.5)
15	484	9.42%	148	6.19%	3.0	< 1E-5	(2.2, 4.1)
16	308	5.99%	111	4.64%	3.0	< 1E-5	(2.0, 4.4)
17	190	3.70%	68	2.84%	2.9	3.0E-04	(1.7, 4.8)
18–32	214	4.17%	78	3.26%	2.7	5.2E-03	(1.4, 5.4)

Abbreviations. OR: odds ratio; EMP P: empirical p-value.