Two Genome-wide Association Studies of Aggressive Prostate Cancer Implicate Putative Prostate Tumor Suppressor Gene *DAB2IP*

David Duggan, Siqun L. Zheng, Michele Knowlton, Debbie Benitez, Latchezar Dimitrov, Fredrik Wiklund, Christiane Robbins, Sarah D. Isaacs, Yu Cheng, Ge Li, Jielin Sun, Bao-Li Chang, Leslie Marovich, Kathleen E. Wiley, Katarina Bälter, Pär Stattin, Hans-Olov Adami, Marta Gielzak, Guifang Yan, Jurga Sauvageot, Wennuan Liu, Jin Woo Kim, Eugene R. Bleecker, Deborah A. Meyers, Bruce J. Trock, Alan W. Partin, Patrick C. Walsh, William B. Isaacs, Henrik Grönberg, Jianfeng Xu, John D. Carpten

Background The consistent finding of a genetic susceptibility to prostate cancer suggests that there are germline sequence variants predisposing individuals to this disease. These variants could be useful in screening and treatment.

Methods We performed an exploratory genome-wide association scan in 498 men with aggressive prostate cancer and 494 control subjects selected from a population-based case–control study in Sweden. We combined the results of this scan with those for aggressive prostate cancer from the publicly available Cancer Genetic Markers of Susceptibility (CGEMS) Study. Single-nucleotide polymorphisms (SNPs) that showed statistically significant associations with the risk of aggressive prostate cancer based on two-sided allele tests were tested for their association with aggressive prostate cancer in two independent study populations composed of individuals of European or African American descent using one-sided tests and the genetic model (dominant or additive) associated with the lowest value in the exploratory study.

Results Among the approximately 60 000 SNPs that were common to our study and CGEMS, we identified seven that had a similar (positive or negative) and statistically significant ($P<.01$) association with the risk of aggressive prostate cancer in both studies. Analysis of the distribution of these SNPs among 1032 prostate cancer patients and 571 control subjects of European descent indicated that one, rs1571801, located in the *DAB2IP* gene, which encodes a novel Ras GTPase-activating protein and putative prostate tumor suppressor, was associated with aggressive prostate cancer (one-sided P value = .004). The association was also statistically significant in an African American study population that included 210 prostate cancer patients and 346 control subjects (one-sided P value = .02).

Conclusion A genetic variant in *DAB2IP* may be associated with the risk of aggressive prostate cancer and should be evaluated further.

J Natl Cancer Inst 2007;99:1836–44

The consistent finding of a genetic susceptibility to prostate cancer suggests that there are germline sequence variants that predispose individuals who carry them to the disease (1). Single-nucleotide polymorphisms (SNPs) may be causally related to prostate cancer risk or be indirectly associated with prostate cancer risk through linkage disequilibrium with a causal sequence variant. Risk-associated SNPs will have different frequencies among men with or without prostate cancer and can be detected using genetic association studies. By testing associations of disease with a large number of SNPs that capture most of the genetic variation in a population, genome-wide association studies can identify variants

Affiliations of authors: Divisions of Genetic Basis of Human Disease (DD, MK, DB, LM) and Division of Integrative Cancer Genomics (CR, JC), Translational Genomics Research Institute, Phoenix, AZ; Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, NC (SZ, LD, YC, GL, JS, BC, WL, JWK, ERB, DAM, JX); Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (FW, KB, HOA, HG); Department of Urology, Johns Hopkins Medical Institutions, Baltimore, MD (SDI, KEW, MG, GY, JS, BJT, AWP, PCW, WBI); Department of Surgical and Perioperative Sciences, Urology and Andrology, Umeå University Hospital, Umeå, Sweden (PS); Department of Epidemiology, Harvard University, Boston, MA (HOA).

Correspondence to: Jianfeng Xu, MD, DrPH, Center for Human Genomics, Medical Center Blvd, Winston-Salem, NC 27157 (e-mail: jxu@wfubmc.edu), Henrik Grönberg, MD, PhD, Karolinska Institutet, Nobels väg 12 A, PO Box 281, Stockholm SE-171 77, Sweden (e-mail: henrik.gronberg@ki.se), or William B. Isaacs, PhD, Marburg 115, Johns Hopkins Hospital, 600 North Wolfe Street, Baltimore, MD 21287 (e-mail: wisaacs@jhmi.edu).

See "Funding" and "Notes" following "References."

DOI: 10.1093/jnci/djm250

that are associated with disease risk. For example, two genome-wide association studies of prostate cancer in Iceland (2) and the United States (3) have recently identified novel prostate cancer risk variants at the chromosomal locus 8q24.

In this study, we sought to identify SNPs that are associated with aggressive prostate cancer by performing a genome-wide association study that compared subjects with aggressive prostate cancer (defined by stage and/or grade) and those without prostate cancer. In the exploratory phase of our study, we combined the data from our study population, which was from a population-based case–control study in Sweden, with the publicly available results for aggressive prostate cancer from the United States Cancer Genetic Markers of Susceptibility (CGEMS) Study (http://cgems.cancer.gov/data). The focus on aggressive disease has at least two potential advantages. Aggressive prostate cancers are more likely to have a poor outcome, and, thus, identifying genetic variants that predict risk of aggressive disease could have an impact on decision making related to prostate cancer screening, diagnosis, and treatment. Furthermore, focusing the analysis on aggressive prostate cancer, a phenotype that is less common and easily distinguishable from unaffected men, may reduce misclassification of case subjects and control subjects and therefore increase the power to detect associations. For SNPs that were associated with aggressive prostate cancer in both genome-wide association analyses, we performed a confirmation study in two independent populations of patients with aggressive prostate cancer.

CONTEXT AND CAVEATS

Prior knowledge

The genetic variants that may predispose individuals to prostate cancer are largely unknown.

Study design

Single-nucleotide polymorphisms (SNPs) were tested for their association with aggressive prostate cancer in a population of case subjects and control subjects. SNPs that appeared to be strongly associated with the risk of aggressive prostate cancer in the exploratory analyses were then further tested in two independent populations.

Contribution

This study found evidence that a putative tumor suppressor gene may be associated with the risk of aggressive prostate cancer.

Implications

If confirmed, this work and previous studies that have associated genetic polymorphisms with prostate cancer may furnish a basis for improved screening and treatment.

Limitations

Possible effects of population stratification could not be fully controlled for in this study.

Subjects and Methods

Study Samples

Subjects in the Two Genome-wide Association Studies of Aggressive Prostate Cancer. Results from two genome-wide association studies of aggressive prostate cancer were combined to identify SNPs that are associated with aggressive prostate cancer. In the first study, we selected 498 subjects with aggressive prostate cancer and 494 control subjects that matched the age distribution of case subjects from CAPS (Cancer of the Prostate in Sweden), a population-based case–control study in Sweden (4). The 498 patients with aggressive prostate cancer met at least one of the following criteria based on the biopsy specimen: clinical stage T3/T4, N+, M+, differential grade III, Gleason score of 8 or higher, or preoperative serum PSA of at least 50 ng/mL. Differential grade is a World Health Organization grading system that is commonly used in Sweden (5). In the second genome-wide association study, we analyzed data from all 737 patients with aggressive prostate cancer, defined as clinical stage T3/T4 or Gleason score of 7 or higher based on biopsy specimens, and 1105 age-matched control subjects from the National Cancer Institute CGEMS Initiative (cgems.cancer.gov) (3). All of the subjects selected from CGEMS were of European ancestry.

Subjects in the Confirmation Study of Aggressive Prostate Cancer. Patients with aggressive prostate cancer (n = 1242) and unaffected control subjects (n = 917) from Johns Hopkins Hospital were studied to confirm the findings from the two genome-wide association studies. Prostate cancer patients with aggressive prostate cancer were selected from among patients undergoing radical prostatectomy for treatment of prostate cancer at Johns Hopkins Hospital between 1999 and 2006. Tumors from each patient were graded and staged using uniform criteria (6). Based on radical prostatectomy specimens, patients that met at least one of the following criteria were classified as having aggressive disease: pathologic stage T3/T4, N+, M+, Gleason score of 7 or more, or preoperative serum PSA of at least 20 ng/mL. During the same time period, men undergoing screening for prostate cancer at the Johns Hopkins Hospital and other sites in the greater Baltimore metropolitan area were asked to participate as control subjects in the prostate cancer study. Serum PSA levels, results of digital rectal examination, and demographic information were available for these control subjects. The Johns Hopkins Hospital study included cohorts of both European Americans (JHH-EA), with 1032 subjects with aggressive prostate cancer and 571 unaffected control subjects, and African Americans (JHH-AA), with 210 subjects with aggressive prostate cancer and 346 unaffected control subjects. The study was approved by the institutional review board and is consistent with Health Insurance Portability and Accountability Act of 1996.

Additional Prostate Cancer Patients in the Cancer Genetic Markers of Susceptibility and Johns Hopkins Hospital Study Populations. We evaluated the allele frequency of one SNP, rs1571801, among additional patients with nonaggressive disease in the CGEMS and the Johns Hopkins Hospital studies. There were 624 subjects with nonaggressive disease in the CGEMS study (nonaggressive was defined as clinical Gleason score < 7 and stage < III). In the study performed at Johns Hopkins Hospital, we enrolled an additional 528 patients of European descent and 156 patients of African American descent who were treated at this institution between 1999 and 2006 and had pathologic Gleason score of less than 7 and organ-confined disease.

Investigations were approved by respective local institutional review boards in accordance with assurances filed with and approved by the US Department of Health and Human Services.

DNA Samples

DNA samples isolated from whole blood were available for all subjects in CAPS and for the Johns Hopkins Hospital control subjects. DNA samples for the Johns Hopkins Hospital prostate cancer patients were isolated from frozen seminal vesicle tissues. From a subset of these subjects (n = 20) in which DNA from both blood and seminal vesicle tissues was available, we observed 100% concordance in genotypes for 162 SNPs chosen randomly across the genome (data not shown).

Genotyping for the Genome-wide Association Studies

For the genome-wide association study in CAPS, the genotyping was performed at Translational Genomics Research Institute, Phoenix, AZ, using the GeneChip Human Mapping 500-K Array Set from Affymetrix (Santa Clara, CA). GeneChips were processed according to the manufacturer's recommendations. Genotyping of 30-K genomic fill-in SNPs and 20-K nonsynonymous SNPs was performed at Affymetrix using Molecular Inversion Probe technology. The 30-K fill-in SNPs include SNPs in the gaps of the Affymetrix 500-K SNPs, and their inclusion improved the genome-wide coverage by approximately 7% (data not shown). The 20-K nonsynonymous SNPs primarily represent validated SNPs that result in an amino acid change from the SNP database (dbSNP). To ensure the quality of the genotype data, automated liquid handling systems (Biomek FX, Beckman Coulter, Fullerton, CA) were used whenever possible, for example, when aliquoting DNA samples and setting up 500-K assays. Two quality control DNA samples (one from the Centre d'Etude du Polymorphisme Humain and one supplied by Affymetrix) were included in every 96-well plate, and their genotypes were compared with the genotypes as determined by HapMap and Affymetrix, respectively. The Bayesian Robust Linear Model with Mahalanobis distance classifier (BRLMM) algorithm from Affymetrix was used to make genotype calls, and the average genotyping call rate (i.e., number of SNPs being called by BRLMM algorithm/total number of SNPs) was 99.1%. Genotype concordance for the positive controls was greater than 99%. HumanHap300 BeadChips from Illumina (San Diego, CA) were used for the genome-wide association study in CGEMS (3).

Genotyping for the Confirmation Study

SNPs that were statistically significantly associated with advanced prostate cancer in both genome-wide association studies and SNPs used for mapping were genotyped in aggressive prostate cancer patients and unaffected control subjects of the Johns Hopkins Hospital Study using the MassArray System from Sequenom (San Diego, CA). Polymerase chain reaction (PCR) and extension primers for these SNPs were designed using the MassARRAY Assay Design 3.0 software (Sequenom). PCR and extension reactions were performed according to the manufacturer's instructions, and extension product sizes were determined by mass spectrometry using the Sequenom iPLEX system. Duplicated and water samples, to which the technician was blind, were included in each 96-well plate as PCR negative controls. The genotype call rates of these SNPs were more than 98% and the average concordance rate between samples was more than 99%. As another quality control check, we tested the frequency of each SNP for agreement with Hardy–Weinberg equilibrium.

Ancestral Informative Microsatellite Markers

As previously described (6), the following microsatellite markers were typed in subjects of JHH-AA to estimate the genetic ancestry: D1S2630, D1S2847, D1S466, D1S493, D2S166, D3S1583, D3S4011, D3S4559, D4S2460, D4S3014, D5S1967, DG5S802, D6S1037, D8S1719, D8S1746, D9S1777, D9S1839, D9S2168, D10S1698, D11S1321, D11S4206, D12S1723, D13S152, D14S588, D17S1799, D17S745, D18S464, D19S113, D20S878, and D22S1172. The genotyping method has been described in detail previously (7). Briefly, following multiplex PCR using fluorescently labeled primers, the resulting PCR fragments were separated using capillary electrophoresis by an ABI 3730 sequencer (Applied Biosystems, Foster City, CA).

Statistical Methods

A Hardy–Weinberg equilibrium test was performed for each SNP in case subjects or control subjects of CAPS, using the Fisher's exact test. SNPs for which there was statistically significant deviation from Hardy–Weinberg equilibrium ($P<.05$) in both case subjects and control subjects or for which there was statistically significant deviation from Hardy–Weinberg equilibrium ($P<.01$) in control subjects only were not analyzed further. The Haploview program (http://www.broad.mit.edu/mpg/haploview) was used to estimate linkage disequilibrium and perform haplotype block analysis (8). We performed association tests for each of the remaining SNPs using an allele test; the allele frequency of each SNP in case subjects and control subjects was compared using a two-sided chi-square test. Limiting the association tests to allele tests decreases the number of tests but may miss SNPs that are associated with prostate cancer risk under some genetic models such as an over-dominant mode of inheritance. We focused on the SNPs that were genotyped in both genome-wide association studies and selected SNPs that were statistically significant at a P value of less than .01 on the basis of the allele test and had the same direction of association in both genome-wide association studies. For these SNPs (n = 7), we also performed a series of association tests using various genetic models for the risk allele (dominant, recessive, additive, and two-degree-of-freedom general models) to identify the model corresponding to the smallest P value. To confirm the associations of these seven SNPs with aggressive prostate cancer, we compared the genotype frequencies of these SNPs among patients with aggressive prostate cancer and unaffected control subjects of the JHH-EA and JHH-AA cohorts using the best genetic model obtained in the combined genome-wide association studies using a one-sided test. A one-sided test was used because this was a confirmation study, and only the associations in the same direction would be considered as confirmed. When the SNPs were confirmed, various genetic models, including additive, dominant, or recessive, were tested using unconditional logistic regression and adjusted for age and, in the JHH-AA cohort only, individual ancestry proportion. Odds ratios (ORs) and 95% confidence intervals (CIs) for prostate cancer were estimated for men having risk genotypes compared with men

having nonrisk genotypes under these genetic models. The dependent variable in the logistic regression analysis was aggressive prostate cancer when we compared patients with aggressive prostate cancer with control subjects and nonaggressive prostate cancer when we compared patients with nonaggressive prostate cancer with control subjects. Finally, the results from multiple case–control populations were combined using a Mantel–Haenszel model (9) in which the populations were allowed to have different population frequencies for alleles but were assumed to have a common odds ratio. The homogeneity of odds ratios among different study populations was tested using Breslow–Day chi-square test.

Associations between haplotypes of SNPs and prostate cancer risk were performed using a score test developed by Schaid et al. (10), as implemented in the computer program Haplo.stat (http://www.mayo.edu/statgen).

Test for Population Stratification

Thirty unlinked microsatellite markers selected as ancestry-informative markers were genotyped among subjects in the JHH-AA cohort (7). We used the Structure program (http://pritch.bsd.uchicago.edu/software/structure2_2.html) to infer the number of ancestral populations and to estimate the proportion of ancestry for each individual (11). The individual ancestry proportion was used as a covariate in the association tests to minimize the effect of potential population stratification.

Assessment for Potential Systematic Bias in Genome-wide Association Tests in the Cancer of the Prostate in Sweden Study

Differences in genetic background between case subjects and control subjects (i.e., population stratification) may lead to a bias toward false positive associations in the results of genome-wide association tests (12). We tested for this potential bias by comparing the observed and expected numbers of statistically significant SNPs under a null hypothesis of no disease association using a Kolmogorov–Smirnov test and a quantile–quantile plot. A large number of independent SNPs that are not associated with prostate cancer risk are needed for this test. Although it is impossible to know which SNPs are not associated with prostate cancer a priori, we approximated the condition of independence (from disease risk) by randomly selecting 1 SNP per Mb throughout the genome. To remove potential linkage disequilibrium among these SNPs, we estimated the pairwise linkage disequilibrium and obtained a set of SNPs (n = 2356) that had pairwise $r^2 = 0$ among them. We performed a single SNP association test for each of these SNPs using an unconditional logistic regression model that assumed an additive mode of inheritance. The distribution of the observed χ^2 from these SNPs was compared with that under a null hypothesis of no disease association. The D statistic of the Kolmogorov–Smirnov test was 0.026 ($P = .08$), suggesting no systematic bias in the association results. The quantile–quantile plot of these data is presented in Supplementary Fig. 1 (available online). The slope of the observed χ^2 was 1.01; thus, there was no evidence for a systematic upward bias (12).

Results

Among the 60 275 SNPs that were genotyped in both our study and CGEMS, 81 SNPs were statistically significantly ($P<.05$ from two-sided allele tests) associated with aggressive prostate cancer and with consistent effects on risk (positive or negative) in both genome-wide association studies (Supplementary Table 1, available online). Eleven SNPs had a statistically significant association with prostate cancer using a threshold of $P<.01$ (two-sided test) in both studies (Table 1), seven of which had consistent risk associations in both studies. Data from the two genome-wide association studies were combined, and a series of association tests using various genetic models (dominant, recessive, additive, and two-degree-of-freedom general models) were performed to identify the best model

Table 1. SNPs that were associated with prostate cancer risk in the CAPS and CGEMS studies*

				CAPS			CGEMS						
				MAF			MAF			Direction of association	CAPS + CGEMS		JHH
Chr	Position	SNPs	MA	Agg.	Cont.	P†	Agg.	Cont.	P†		P‡	Model	P§
3	45,980,482	rs1545985	G	0.401	0.322	.0003	0.400	0.349	.0024	Same	6.61×10^{-6}	add	.31
3	45,937756	rs7652331	T	0.354	0.299	.009	0.381	0.327	.0011	Same	4.47×10^{-5}	dom	.11
4	56,983,965	rs629242	T	0.230	0.167	.0005	0.256	0.201	.0001	Same	7.25×10^{-7}	add	.12
4	147,108,571	rs13149290	T	0.189	0.237	.0098	0.201	0.239	.0093	Same	2.56×10^{-5}	dom	.04
5	141,007,068	rs251177	C	0.226	0.282	.0046	0.221	0.260	.0077	Same	.000188	add	.37
9	121,506,927	rs1571801	A	0.308	0.246	.0019	0.281	0.241	.0076	Same	2.84×10^{-5}	dom	.004‖
13	50,722,329	rs10492519	G	0.476	0.418	.0099	0.472	0.415	.0007	Same	5.6×10^{-6}	dom	.42
4	141,181,340	rs736349	T	0.063	0.100	.0026	0.098	0.073	.0093	Opposite			
5	97,449232	rs257226	A	0.463	0.537	.001	0.513	0.466	.0066	Opposite			
9	121,031239	rs942152	T	0.498	0.557	.0091	0.419	0.471	.0025	Opposite			
20	23,092,070	rs6083025	G	0.421	0.485	.0047	0.415	0.372	.0098	Opposite			

* SNPs = single nucleotide polymorphisms; CAPS = Cancer of the Prostate in Sweden (4); CGEMS = Cancer Genetic Markers of Susceptibility; JHH = Johns Hopkins Hospital; Chr = chromosome; MA = minor allele; MAF = minor allele frequency; Agg. = patients with aggressive prostate cancer; Cont. = control subjects; add = additive; dom = dominant.

† Test for allele frequency differences using a two-sided chi-square test.

‡ Test for genotype frequency differences using a Wald chi-square test (two-sided), adjusted for age.

§ Test for genotype frequency differences using a Wald chi-square test (one-sided) assuming the best fitted genetic model obtained from the exploratory analyses.

‖ Statistically significant after adjustment for seven independent tests.

Table 2. Association of rs1571801 and prostate cancer risk according to study group and prostate cancer aggressiveness*

Study group	Genotype	Number (%) of subjects		OR (95% CI)	*P*†
		Case subjects	Control subjects		
		Aggressive disease			
CAPS	CC	233 (46.8)	281 (56.9)	1.00 (referent)	
	CA	223 (44.8)	183 (37.0)		
	AA	42 (8.43)	30 (6.07)		
	CA/AA	265 (53.2)	213 (43.1)	1.50 (1.17 to 1.93)	.0015
CGEMS	CC	347 (51.6)	626 (58.0)	1.00	
	CA	272 (40.5)	388 (35.9)		
	AA	53 (7.89)	66 (6.11)		
	CA/AA	325 (48.4)	454 (42.0)	1.29 (1.06 to 1.57)	.0096
JHH-EA	CC	553 (53.6)	345 (60.4)	1.00	
	CA	409 (39.6)	186 (32.6)		
	AA	70 (6.78)	40 (7.01)		
	CA/AA	479 (46.4)	226 (39.1)	1.32 (1.07 to 1.63)	.0083
JHH-AA	CC	144 (69.2)	267 (77.2)	1.00	
	CA	59 (28.4)	69 (19.9)		
	AA	5 (2.40)	10 (2.89)		
	CA/AA	64 (30.8)	79 (22.8)	1.50 (1.02 to 2.21)	.039
Combined (JHH-EA/ JHH-AA)	CA/AA vs CC			1.36 (1.13 to 1.63)	.001‡
		Non-aggressive disease			
CGEMS	CC	242 (50.6)	626 (58.0)	1.00	
	CA	195 (40.8)	388 (35.9)		
	AA	41 (8.58)	66 (6.11)		
	CA/AA	236 (49.4)	454 (42.0)	1.34 (1.08 to 1.67)	.007
JHH-EA	CC	295 (56.1)	345 (60.4)	1.00	
	CA	193 (36.7)	186 (32.6)		
	AA	38 (7.22)	40 (7.01)		
	CA/AA	231 (43.9)	226 (39.1)	1.20 (0.94 to 1.52)	.14
Combined (CGEMS/ JHH-EA)	CA/AA vs CC			1.28 (1.09 to 1.50)	.003‡
JHH-AA	CC	113 (72.9)	267 (77.2)	1.00	
	CA	38 (24.5)	69 (19.9)		
	AA	4 (2.58)	10 (2.89)		
	CA/AA	42 (27.1)	79 (22.8)	1.26 (0.81 to 1.94)	.3
Combined (CGEMS/ JHH-EA/JHH-AA)	CA/AA vs CC			1.27 (1.10 to 1.48)	.0017‡

* OR = odd ratio; CI = confidence interval; CAPS = Cancer of the Prostate in Sweden (4); CGEMS = Cancer Genetic Markers of Susceptibility; JHH = Johns Hopkins Hospital; JHH-EA = European American subjects in JHH; JHH-AA: African American subjects in JHH.

† Based on Wald chi-square test (two-sided), assuming a dominant model and adjusted for age.

‡ *P* values are based on the Mantel–Haenszel test.

(i.e., the model for which the *P* value was smallest) for each of the SNPs (data not shown). We then evaluated these seven SNPs among 1032 patients with aggressive prostate cancer and 571 unaffected control subjects from the JHH-EA study population using the best model obtained in the combined genome-wide association studies for each SNP. We found an association for one of the seven SNPs (rs1571801 at 9q33, $P = .004$ under a dominant model for allele A, one-sided test). The association was statistically significant at a 5% type I error level after adjusting for seven tests using a Bonferroni correction ($P_{adjusted} = .028$). The association was even stronger when the analysis was limited to the 578 case subjects with Gleason score greater than 8, $P = .001$ (one-sided test). The frequency of allele A of the SNP rs1571801 was consistently higher among case subjects than among control subjects in the three populations (frequency in case subjects/frequency in control subjects = 0.31/0.25, 0.28/0.24, and 0.27/0.23 in CAPS, CGEMS, and JHH-EA, respectively). To further confirm this association and test it among populations of non-European ancestry, we genotyped rs1571801 among 210 patients with aggressive prostate cancer and 346 unaffected control subjects from the JHH-AA study population. Under a dominant model for allele A, we found that the SNP was statistically significantly associated with aggressive prostate cancer ($P = .02$, one-sided test).

As shown in Table 2, compared with men who have CC genotype at rs1571801, men who carried risk genotypes (CA and AA) had a statistically significantly increased risk of aggressive prostate cancer in each of the four study populations. The estimated ORs in the discovery study populations were 1.50 (95% CI = 1.17 to 1.93, $P = .0015$, two-sided test) in CAPS and 1.29 (95% CI = 1.06 to 1.57, $P = .0096$, two-sided test) in CGEMS. The combined OR in the two independent confirmation study populations (JHH-EA and JHH-AA) was 1.36 (95% CI = 1.13 to 1.63, $P = .001$, two-sided test).

To test the possibility that the SNP rs1571801 is associated with risk of overall prostate cancer rather than risk of aggressive prostate cancer alone, we estimated the allele frequency of rs1571801 among patients with nonaggressive prostate cancer in several studies, including 624 patients from CGEMS and 528 patients of European descent and 156 patients of African American descent in the Johns Hopkins Hospital Study. The allele frequency in patients of European ancestry with nonaggressive prostate cancer was 0.29 in CGEMS and 0.25 in the JHH-EA population, similar to that in patients with aggressive prostate cancer. When the patients in the CGEMS and JHH-EA populations were combined, carriers of the allele A of rs1571801 had a statistically significantly increased risk of nonaggressive prostate cancer (OR = 1.28, 95% CI = 1.09 to 1.50, P = .003, two-sided test). No heterogeneity was detected in the estimates of odds ratios among the two studies (P = .48 based on a Breslow–Day chi-square test). By contrast, the frequency of allele A of rs1571801 in nonaggressive patients in the JHH-AA cohort was 0.15, in between that of aggressive prostate cancer (0.17) and control subjects (0.13). Carriers of the risk allele A had an increased risk of nonaggressive prostate cancer, but the increase was not statistically significant (OR = 1.26, 95% CI = 0.81 to 1.94, P = .3, two-sided test). When subjects with nonaggressive prostate cancer and control subjects from the CGEMS study and the JHH-EA and JHH-AA populations were combined, carriers of allele A had a moderately increased risk for nonaggressive prostate cancer (OR =1.27, 95% CI = 1.10 to 1.48; P = .0017, two-sided test).

To fine map the genomic region of the SNP rs1571801, which is at 121,506,927 bp on chromosome 9 (NCBI Build 35), we examined the association results for all of the surrounding SNPs in the genome-wide association studies of CAPS and CGEMS (including more than 300 000 SNPs in the initial phase and 240-K SNPs of the second phase of CGEMS). A 19-kb region, flanked by SNPs rs7025486 at 121,501,957 bp (proximal) and rs7047340 at 121,520,958 bp (distal) was studied. Within the region, we selected all of the independent polymorphic SNPs in the CEU (Utah residents with ancestry from northern and western Europe) population of the HapMap database that had pairwise r^2 greater than 0.8 (n = 11) and all reported SNPs in the dbSNPs (n = 15). Genotyping these SNPs among a subset of the patients with aggressive prostate cancer in the JHH-EU cohort (n = 578) and control subjects (n = 571) revealed that all 15 dbSNPs were monomorphic and that none of the 11 HapMap SNPs was statistically significantly associated with risk of aggressive prostate cancer (Supplementary Table 2, available online). However, we found that the approximately 6-kb region immediately flanking rs1571801 did not contain any known polymorphic SNPs (between rs2150711 at 121,505,429 and rs10119920 at 121,511,808). Therefore, we attempted to identify novel SNPs in this region among 48 JHU-EA prostate cancer case subjects and 48 control subjects by sequencing. This sample size had power greater than 99% to detect variants with minor allele frequency greater than 5%, as estimated by Kruglyak and Nickerson (13). Analysis of the sequencing data revealed four novel sequence variants (gasp66154, gasp66601, gasp68276, and gasp68492), all of which mapped to a 2.4-kb region that includes rs1571801. We genotyped these four SNPs; two of these novel SNPs, one proximal (gasp66601 at 121,506,334) and one distal (gasp68492 at 121,508,225) to rs1571801, were statistically significantly associated with risk of prostate cancer (Fig. 1, A; Supplementary Table 2, available online). These two SNPs are in strong linkage disequilibrium with rs1571801 (r^2 = 0.98; Fig. 1, B). The remaining two SNPs (gasp68276 and gasp66154), both of which are in low linkage disequilibrium with rs1571801 (r^2 = 0.02), were not statistically significantly associated with prostate cancer risk. A haplotype analysis of these four SNPs together with rs1571801 provided evidence for association with aggressive prostate cancer risk (P = .02, two-sided test) and revealed one haplotype with frequencies of 0.26 and 0.23 in case subjects and control subjects, respectively. The results of association tests for the haplotype of a broader region including all eight SNPs in block 1 (Fig. 1, B) were not statistically significant, suggesting the prostate cancer–associated sequence variants at this region are most likely near rs1571801. The localization of prostate cancer association at this region to rs1571801 is further supported by the results from JHH-AA cohort in which rs1571801 was statistically significantly associated with prostate cancer risk, but the two flanking SNPs, gasp66601 and gasp68429, were not (P = .19 and .18, respectively, two-sided test).

Discussion

The power of genome-wide association studies to systematically and objectively detect disease genetic variants associated with disease risk was demonstrated by two recent studies that independently revealed the association of 8q24 variants and risk of prostate cancer (2,3). The 8q24 risk variants were among the SNPs most strongly associated with prostate cancer risk in both studies; $P = 1.6 \times 10^{-14}$ and 9.7×10^{-5} for the SNP rs1447295 in the Icelandic population (2) and the CGEMS study, respectively (3). In this study, we utilized a combined approach to search for variants that increase risk of aggressive prostate cancer, the most deadly form of the disease.

Among approximately 60 000 SNPs that were studied in the genome-wide association studies of aggressive prostate cancer in Sweden (CAPS) and in the CGEMS study, we identified seven SNPs that were statistically significantly associated with aggressive prostate cancer risk at a threshold of P greater than .01 and had a similar association with risk in both studies. Association with aggressive prostate cancer was confirmed for one of the seven SNPs in two independent study populations from Johns Hopkins Hospital, one composed of Europeans and the other of African Americans. Carriers of risk allele had a statistically significantly increased risk of aggressive prostate cancer in these two independent confirmation populations, with combined OR = 1.36, 95% CI = 1.13 to 1.63, P = .001. The risk allele was also associated with a statistically significant but slightly lower risk for nonaggressive prostate cancer (OR = 1.27, 95% CI = 1.10 to 1.48, P = .0017).

SNP rs1571801 maps within the *DAB2IP* gene (MIM 190020). There are two *DAB2IP* isoforms annotated in Entrez Gene, isoforms 1 and 2, and a third larger transcript listed in the RefSeq (the National Center for Biotechnology Information Reference Sequence) database. SNP rs1571801 is located in intron 1 of this third transcript, approximately 14-kb upstream of exon 2, and approximately 34-kb upstream of the transcription start site of Entrez isoform 1.

DAB2IP encodes a novel Ras GTPase-activating protein that is a potent negative regulator of RAS/PKC-elicited signal

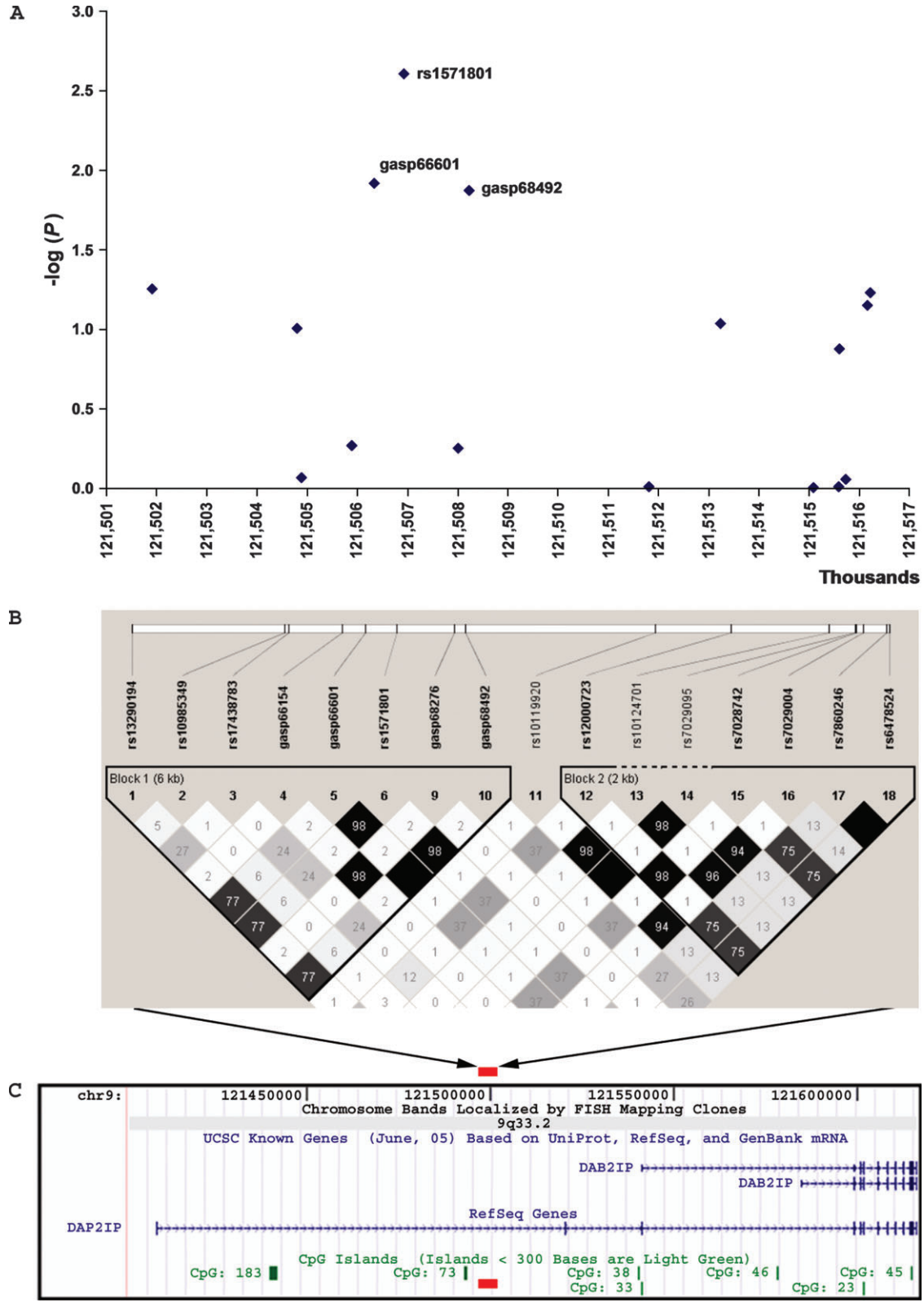

Fig. 1. A schematic view of single-nucleotide polymorphisms (SNPs) in the 16-kb flanking region of rs1571801 and genomic region of the *DAB2IP* gene. (**A**) Association tests for 16 SNPs at 9q33.2 and prostate cancer among 1032 aggressive prostate cancer case subjects and 571 control subjects, all of which are European Americans, collected at Johns Hopkins Hospital (JHH-EA). The y-axis is the $-\log_{10}(P)$ for association tests under a rare allele dominant model. The x-axis is the physical position on chromosome 9 based on UCSC Build 35. Each **dot** represents a SNP. (**B**) Pairwise r^2 between the 16 SNPs among 1032 case subjects and 571 control subjects from JHH-EA. Two haplotype blocks, estimated using the default parameters of Haploview, are shown. (**C**) Genomic view at the *DAB2IP* gene and the location of CpG Islands at the region, as obtained from UCSC. The 16-kb flanking region of rs1571801 is represented by the **red bar**.

transduction (14). *DAB2IP* is expressed in normal prostate epithelial cells (15), but the expression is typically decreased in prostate cancer cells (16), and induced expression by DNA transfection suppresses the growth of prostate cancer cells. Recently, the reduced expression of *DAB2IP* has been shown to correlate with the increased expression of the transcriptional repressor *EZH2* (17), a component of the polycomb complex 2/3. The increased expression of *EZH2* has been shown to be one of the most robust markers of aggressive prostate cancer (18). The association of polycomb repressor complexes with the *DAB2IP* promoter in prostate cancer cells but not normal epithelium is consistent with the findings of frequent promoter hypermethylation-associated inactivation in prostate and other cancers, including breast cancers (19,20). The DAB2IP protein also binds to and thereby activates MAP3K5 (also referred to as apoptosis signal regulating kinase 1) and is required for tumor necrosis factor–mediated Jun kinase signaling-mediated cell apoptosis (21,22). Finally, a *DAB2IP*-inactivating gene fusion event has been identified in acute myelogenous leukemia (23). Thus, considerable circumstantial evidence and the diverse functional properties of the protein encoded by this gene make it an attractive candidate for a prostate cancer aggressiveness risk gene.

Although the overall statistical evidence for the association between rs1571801 and prostate cancer risk is strong, and the association is supported by its location in a tumor suppressor gene, care should be exercised when interpreting our findings. This SNP was selected from screening 60 275 SNPs in the genome from two genome-wide association studies, with $P = .0019$ in the CAPS study and $P = .0076$ in the CGEMS study from allele tests. It is likely ($P = .0019 \times .0076 \times .5 \times 60275 = .37$) that we would identify two such P values with the same (positive or negative) qualitative association with risk under the null hypothesis of no association. The primary evidence for its association with aggressive prostate cancer comes from the confirmation in the JHH-EA cohort (nominal one-sided P value = .004) after testing for associations with seven SNPs and in the smaller JHH-AA cohort (nominal one-sided $P = .02$ after a single test). Careful examination of the frequency of allele A in the case subjects (0.27) and control subjects (0.23) of the JHH-EA indicated the statistically significant difference is mainly due to the low allele frequency among the control subjects in this population. Based on 1145 breast cancer case subjects (0.26) and 1142 control subjects (0.26) of the CGEMS breast cancer study, the best estimate of the population frequency of allele A for the SNP rs1571801 in the JHH-EA is approximately 0.26. However, the lower estimate of frequency of allele A for the SNP in the control subjects of JHH-EA could be due to the fact that the control subjects were screened for prostate cancer, lowering the number of subjects with aggressive prostate cancer in this group. The lower allele frequency observed in the control subjects of the CGEMS prostate cancer study could be a result of extensive disease screening as well. Finally, although the individual ancestry proportion was adjusted in the association tests in JHH-AA population, no such correction was made in the JHH-EA population. Therefore, the statistically significant association of rs1571801 with the risk of advanced prostate cancer could be an artifact of population stratification.

In summary, we report the discovery of a statistically significant association of prostate cancer risk with an allele in the *DAB2IP* gene, a prostate cancer tumor suppressor. Our study is among the first to report the presence of a potentially important prostate cancer aggressiveness locus based upon genome-wide association analysis. However, we cannot rule out about the possibility of false-positive association. This report is intended to stimulate the conduct of additional confirmation studies for a gene that has strong initial statistical support and biologic relevance as a tumor suppressor gene.

References

(1) Isaacs WB, Xu J, Walsh PC. Hereditary prostate cancer. In: Chung LWK, Isaacs WB, Simons JW, editors. Prostate cancer, biology, genetics, and the new therapeutics. Totowa (NJ): Humana Press; 2001. p. 13–28.

(2) Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631–7.

(3) Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007;39:645–9.

(4) Zheng SL, Augustsson-Balter K, Chang B, Hedelin M, Li L, Adami HO, et al. Sequence variants of toll-like receptor 4 are associated with prostate cancer risk: results from the Cancer Prostate in Sweden Study. Cancer Res 2004;64:2918–22.

(5) Mostofi FK, Sesterhenn I, Sobin LH, editors. *Histological typing of prostate tumours*. Geneva (Switzerland): World Health Organization; 1982. p. 15–23.

(6) Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL; ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma [review]. Am J Surg Pathol 2005;29:1228–42.

(7) Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006;38:652–8.

(8) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5.

(9) Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22:719–48.

(10) Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425–34.

(11) Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. Am J Hum Genet 2000;67:170–81.

(12) Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet 2005;37:1243–6.

(13) Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet 2001;27:234–6.

(14) Wang Z, Tseng CP, Pong RC, Chen H, McConnell JD, Navone N, et al. The mechanism of growth-inhibitory effect of DOC-2/DAB2 in prostate cancer. Characterization of a novel GTPase-activating protein associated with N-terminal domain of DOC-2/DAB2. J Biol Chem 2002;277:12622–31.

(15) Chen H, Pong RC, Wang Z, Hsieh JT. Differential regulation of the human gene DAB2IP in normal and malignant prostatic epithelia: cloning and characterization. Genomics 2002;79:573–81.

(16) Chen H, Toyooka S, Gazdar AF, Hsieh JT. Epigenetic regulation of a novel tumor suppressor gene (hDAB2IP) in prostate cancer cell lines. J Biol Chem 2003;278:3121–30.

(17) Chen H, Tu SW, Hsieh JT. Down-regulation of human DAB2IP gene expression mediated by polycomb Ezh2 complex and histone deacetylase in prostate cancer. J Biol Chem 2005;280:22437–44.

(18) Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 2002;419:624–9.

(19) Yano M, Toyooka S, Tsukuda K, Dote H, Ouchida M, Hanabata T, et al. Aberrant promoter methylation of human DAB2 interactive protein (hDAB2IP) gene in lung cancers. Int J Cancer 2005;113:59–66.

(20) Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, et al. The polycomb group protein EZH2 directly controls DNA methylation. Nature 2006;439:871–4.

(21) Zhang R, He X, Liu W, Lu M, Hsieh JT, Min W. AIP1 mediates TNF-alpha-induced ASK1 activation by facilitating dissociation of ASK1 from its inhibitor 14-3-3. J Clin Invest 2003;111:1933–43.
(22) Zhang H, Zhang R, Luo Y, D'Alessio A, Pober JS, Min W. AIP1/DAB2IP, a novel member of the Ras-GAP family, transduces TRAF2-induced ASK1-JNK activation. J Biol Chem 2004;279:44955–65.
(23) von Bergh AR, Wijers PM, Groot AJ, van Zelderen-Bhola S, Falkenburg JH, Kluin PM, et al. Identification of a novel RAS GTPase-activating protein (RASGAP) gene at 9q34 as an MLL fusion partner in a patient with de novo acute myeloid leukemia. Genes Chromosomes Cancer 2004;39:324–34.

Funding

National Cancer Institute (CA106523, CA105055, and CA95052 to J. X.; CA112517 and CA58236 to W. B. I.; CA86323 to A. W. P.); Department of Defense (PC051264 to J. X.); Swedish Cancer Society (Cancerfonden) (to H. G.); Swedish Research Council (to H. G.). A substantial amount of this research was funded through the Lee and Nancy Hanley Laboratory for Genomic Research.

Notes

D. Duggan and S. Zheng contributed equally to this study.

The authors thank all the study subjects who participated in this study. We also thank the regional cancer registries and the CAPS-steering committee including Drs Jan-Erik Johansson and Eberhart Varenhorst. We acknowledge the support of William T. Gerrard, Mario Duhon, Jennifer and John Chalsty, and David Koch.

The authors declare that they have no competing financial interests.

The authors take full responsibility for the study design, data collection, analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript.

Manuscript received May 9, 2007; revised September 27, 2007; accepted November 2, 2007.