Published in final edited form as:

Nat Genet. 2008 March ; 40(3): 281–283. doi:10.1038/ng.89.

Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer

Julius Gudmundsson^{1,26}, Patrick Sulem^{1,26}, Thorunn Rafnar¹, Jon T Bergthorsson¹, Andrei Manolescu¹, Daniel Gudbjartsson¹, Bjarni A Agnarsson², Asgeir Sigurdsson¹, Kristrun R Benediktsdottir², Thorarinn Blondal¹, Margret Jakobsdottir¹, Simon N Stacey¹, Jelena Kostic¹, Kari T Kristinsson¹, Birgitta Birgisdottir¹, Shyamali Ghosh¹, Droplaug N Magnusdottir¹, Steinunn Thorlacius¹, Gudmar Thorleifsson¹, S Lilly Zheng³, Jieli Sun³, Bao-Li Chang³, J Bradford Elmore⁴, Joan P Breyer⁴, Kate M McReynolds⁴, Kevin M Bradley⁴, Brian L Yaspan⁴, Fredrik Wiklund⁵, Par Stattin⁶, Sara Lindström⁵, Hans-Olov Adami^{5,7}, Shannon K McDonnell⁸, Daniel J Schaid⁸, Julie M Cunningham⁹, Liang Wang⁹, James R Cerhan⁸, Jennifer L St Sauver⁸, Sara D Isaacs¹⁰, Kathleen E Wiley¹⁰, Alan W Partin¹⁰, Patrick C Walsh¹⁰, Sonia Polo¹¹, Manuel Ruiz-Echarri¹², Sebastian Navarrete¹³, Fernando Fuertes¹³, Berta Saez^{14,15}, Javier Godino¹⁴, Philip C Weijerman¹⁶, Dorine W Swinkels¹⁷, Katja K Aben^{18,19}, J Alfred Witjes²⁰, Brian K Suarez²¹, Brian T Helfand²², Michael L Frigge¹, Kristleifur Kristjansson¹, Carole Ober²³, Eiríkur Jonsson²⁴, Gudmundur V Einarsson²⁴, Jianfeng Xu³, Henrik Gronberg⁵, Jeffrey R Smith^{4,25}, Stephen N Thibodeau⁹, William B Isaacs¹⁰, William J Catalona²², Jose I Mayordomo^{12,14,15}, Lambertus A Kiemeny^{18,19,20}, Rosa B Barkardottir², Jeffrey R Gulcher¹, Unnur Thorsteinsdottir¹, Augustine Kong¹, and Kari Stefansson¹

¹deCODE genetics, 101 Reykjavik, Iceland ²Department of Pathology, Landspítali-University Hospital, 101 Reykjavik, Iceland ³Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, USA ⁴Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA ⁵Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, S-17177 Stockholm, Sweden ⁶Department of Urology, Umea University Hospital, S-90185 Umea, Sweden ⁷Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA ⁸Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota 55905, USA ⁹Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA ¹⁰Brady Urological Institute, Johns Hopkins Medical Institution, Baltimore, Maryland 21209, USA ¹¹Division of Radiation Oncology, Quiron Hospital, 50009 Zaragoza, Spain ¹²Division of Medical Oncology, Lozano Blesa University Hospital, University of Zaragoza, 50009 Zaragoza, Spain ¹³Division of Radiation Oncology, Lozano Blesa University Hospital, University of Zaragoza, 50009 Zaragoza, Spain ¹⁴Health Sciences Institute, 50009 Zaragoza, Spain ¹⁵Nanotechnology Institute of Aragon, 50009 Zaragoza, Spain ¹⁶Department of Urology, Rijnstate Hospital, NL-6800 TA Arnhem, The Netherlands ¹⁷Department of Clinical Chemistry, Radboud University Nijmegen Medical Center, 6500 HB Nijmegen, The Netherlands ¹⁸Comprehensive Cancer Center East, 6501 BG Nijmegen, The Netherlands ¹⁹Department of Epidemiology and Biostatistics, Radboud University Nijmegen Medical Center, 6500 HB Nijmegen, The Netherlands

Correspondence should be addressed to K.S. (kstefans@decode.is) or J.G. (julius@decode.is).

²⁶These authors contributed equally to this work.

Note: Supplementary information is available on the Nature Genetics website.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturegenetics/>.

²⁰Department of Urology, Radboud University Nijmegen Medical Center, 6500 HB Nijmegen, The Netherlands ²¹Department of Psychiatry, Washington University School of Medicine, St Louis, Missouri 63110, USA ²²Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611, USA ²³Department of Human Genetics, University of Chicago, Chicago, Illinois 60637, USA ²⁴Department of Urology, Landspítali-University Hospital, 101 Reykjavik, Iceland ²⁵Medical Research Service, VA Tennessee Valley Healthcare System, Nashville, Tennessee 37212, USA

Abstract

We conducted a genome-wide SNP association study on prostate cancer on over 23,000 Icelanders, followed by a replication study including over 15,500 individuals from Europe and the United States. Two newly identified variants were shown to be associated with prostate cancer: rs5945572 on Xp11.22 and rs721048 on 2p15 (odds ratios (OR) = 1.23 and 1.15; $P = 3.9 \times 10^{-13}$ and 7.7×10^{-9} , respectively). The 2p15 variant shows a significantly stronger association with more aggressive, rather than less aggressive, forms of the disease.

Recently, we presented results from a genome-wide association (GWA) analysis of prostate cancer in which we used the Illumina Hap300 SNP chip to genotype Icelandic samples^{1,2}. In subsequent efforts to identify additional variants, we used two criteria to guide our follow-up analysis of the GWA data. First, several studies on prostate cancer risk in relatives of affected men have reported that brothers are at greater risk compared to fathers^{3,4}, thereby implicating an X-linked or recessive mode of inheritance. Second, identifying genetic factors that preferentially associate to aggressive prostate cancer is of utmost importance for improving treatment selection for early stage disease. Hence, we decided to fast track the variant from our GWA analysis (Supplementary Methods online) showing the strongest association among the Illumina Hap300 SNPs located on the X chromosome, as well as a sequence variant on 2p15 that was selected after evaluating variants showing a greater risk in more aggressive prostate cancer (Gleason ≥ 7 and/or T3 or higher and/or node positive and/or metastatic disease) than in less aggressive disease (Gleason < 7 and T2 or lower).

Allele A of the SNP rs5945572 (rs5945572 A) showed the most significant P value of markers located on the X chromosome in a GWA analysis of 23,205 Icelandic samples, with an allelic specific OR of 1.21 ($P = 3.36 \times 10^{-4}$; Table 1). The SNP is located on Xp11.22, a region that has not been previously implicated in prostate cancer and that is correlated with neither the androgen receptor locus on Xq12 nor the HPCX locus on Xq27–28. On chromosome 2p15, allele A of rs2710646 (rs2710646 A) had an OR of 1.16 ($P = 7.79 \times 10^{-4}$; Table 2) for all Icelandic prostate cancer cases and an OR of 1.33 ($P = 3.73 \times 10^{-5}$) for individuals diagnosed with aggressive disease (Supplementary Table 1 online). By examining the Utah CEPH (CEU) HapMap data, we identified numerous SNPs that are substantially correlated with either of the two anchor SNPs, rs5945572 and rs2710646 (D' 0.9 and $r^2 > 0.4$). From this list we selected several SNPs not present on the Illumina Hap300 chip, representing different degrees of correlation with the anchor SNPs, for further genotyping in 1,500 and 800 Icelandic cases and controls, respectively. None of the additional SNPs were found to be more significantly associated with the disease than the anchor SNPs (Supplementary Table 2 online).

We proceeded to genotype rs5945572 and rs2710646 in seven prostate cancer study groups of European descent, with populations from The Netherlands, Spain, Sweden and the United States. However, as the TaqMan assay for rs2710646 on 2p15 failed in design, we replaced it with an assay for a fully correlated SNP, rs721048 (linkage disequilibrium (LD)

characteristics between rs2710646 and rs721048 in Icelanders and in the four HapMap populations (CEU, CHB, JPT and YRI; Supplementary Table 3 online); $D' = 1$; $r^2 = 0.99$), which was used for genotyping in the replication study populations. When results from all seven case-control replication groups were combined, using all prostate cancer cases, they were highly significant for both SNPs with OR = 1.24 ($P = 2.57 \times 10^{-10}$) for rs5945572 A on Xp11, and OR = 1.15 ($P = 2.23 \times 10^{-6}$) for rs721048 A on 2p15 (Tables 1 and 2). When we combined these data with the Icelandic data, the signals at both loci achieved genome-wide significance: rs5945572 A on Xp11.22 had an OR = 1.23 ($P = 3.95 \times 10^{-13}$), and rs721048 A on 2p15 had an OR = 1.15 ($P = 7.66 \times 10^{-9}$) (Tables 1 and 2). Removing all females ($n = 14,135$) from the control group in the analysis of the combined groups yielded essentially identical results for both loci (OR = 1.23 and 1.15 for rs5945572 A and rs721048 A, respectively; Supplementary Methods). In all of the replication groups, the observed effect for the two loci was in the same direction as in the Icelandic samples, except for rs721048 A, which in the Baltimore group showed no effect (OR = 1). However, a test of heterogeneity in the OR of the eight different study groups showed no significant difference for the two loci ($P = 0.89$ and 0.19 for Xp11 and 2p15, respectively). We note that in the results released by the Cancer Genetics Markers of Susceptibility study group⁵ (<http://cgems.cancer.gov/data/>), the two original anchor SNPs, rs5945572 and rs2710646, show nominal, but not genome-wide, significant association to prostate cancer, further supporting the data presented here.

For rs5945572 A on Xp11, the OR seen for cases with younger age at onset (< 65) or aggressive phenotype was the same as for the whole group. The frequency of rs721048 A, on the other hand, was significantly higher among individuals diagnosed with aggressive prostate cancer than among those with less aggressive disease (OR = 1.11; $P = 2.6 \times 10^{-3}$). Comparing the group of individuals with aggressive tumor ($n = 4,787$) to controls gave an allelic OR of 1.22 ($P = 8.7 \times 10^{-10}$) when combined for all study groups (Supplementary Table 1). Specifically, the heterozygous and homozygous carriers of rs721048 A, which are close to 31% and 4% of the general population, have a genotypic OR of 1.22 and 1.49 of being diagnosed with aggressive prostate cancer compared to the noncarriers, respectively. However, rs721048 A, like rs5945572 A, did not show a stronger association in individuals with younger age at onset of prostate cancer.

Both of the variants, rs5945572 on Xp11 and rs2710646 and rs721048 on 2p15, are located within regions characterized by extended linkage disequilibrium, according to the CEU HapMap results. On Xp11.22, the LD block spans the genomic region from 50.832 Mb to 51.553 Mb, covering 721 kb (NCBI Build 35). The 2p15 LD region is even larger—about 1.9 Mb (62.621–64.534 Mb; NCBI Build 35)—and contains several genes. Both of the 2p15 SNPs, rs2710646 and rs721048, are located within one of the introns of the *EHBP1* gene. This gene is thought to be involved in endocytic trafficking and has not previously been implicated in cancer⁶. Through an RT-PCR analysis of various cDNA libraries, we detected expression of *EHBP1* in several different tissue libraries, including those derived from the prostate (data not shown).

On Xp11, several genes have been localized to the 721-kb region of interest, but none have been previously linked with prostate cancer. Possible cancer candidate genes, based on previously described functions, are *GSPT2* and *MAGED1*. *GSPT2* is related to *GSPT1*, a gene encoding a GTP-binding protein that has an essential role at the G1-to S-phase transition of the cell cycle. *MAGED1* has been implicated in programmed cell death through a *JNK*- and *JUN*-dependent mitochondrial pathway^{7,8}. The genes closest to the SNP rs5945572 are *NUDT10* and *NUDT11*, along with a single exon transcript, *LOC340602*, with unknown function (University of California Santa Cruz Genome Browser, May 2004). *NUDT10* and *NUDT11* belong to a subgroup of phosphohydrolases that preferentially

hydrolyze diphosphoinositol polyphosphates (DIPPS)⁹. It has been proposed that members of this protein superfamily may be involved in vesicle trafficking, stress responses, DNA repair and apoptosis^{10,11}. Using RT-PCR analysis of various cDNA libraries, we detected expression of *GSPT2*, *MAGED1*, *LOC340602*, *NUDT10* and *NUDT11* in several different tissue libraries, including those derived from the prostate (data not shown). Which one of these genes, if any, confers the risk identified by the association reported here remains to be shown.

Recently, several sequence variants have been identified that together account for a substantial proportion of the population attributable risk (PAR) of prostate cancer^{1,2,5,12,13}. With the identification of the two variants described here, yet another piece has been added to the puzzle of prostate cancer genetic susceptibility. Both the Xp11 and the 2p15 variants are common and confer a moderate risk, resulting in an estimated PAR of about 7% and 5% in individuals of European descent, respectively. However, the underlying causative biological perturbation linked to these variants remains to be elucidated.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the individuals who participated in the study and whose contribution made this work possible. This project was funded in part by contract number 018827 (Polygene) from the 6th Framework Program of the EU to deCODE genetics, T.R. and L.A.K. The study was supported in part by National Cancer Institute CA105055, CA106523 and CA95052 to J.X., CA112517 and CA58236 to W.B.I., CA86323 to A.W.P., and Department of Defense grant PC051264 to J.X., and in part by a V Foundation award and US Department of Veterans Affairs grants to J.R.S. The principal investigators and corresponding authors for the respective replication study populations are L.A.K. (The Netherlands), J.I.M. (Spain), H.G. (Sweden), W.B.I. (Baltimore), W.J.C. (Chicago), J.R.S. (Nashville) and S.N.T. (Rochester).

References

1. Gudmundsson J, et al. *Nat. Genet.* 2007; 39:631–637. [PubMed: 17401366]
2. Gudmundsson J, et al. *Nat. Genet.* 2007; 39:977–983. [PubMed: 17603485]
3. Narod SA, et al. *Nat. Med.* 1995; 1:99–101. [PubMed: 7585019]
4. Monroe KR, et al. *Nat. Med.* 1995; 1:827–829. [PubMed: 7585188]
5. Yeager M, et al. *Nat. Genet.* 2007; 39:645–649. [PubMed: 17401363]
6. Guilherme A, et al. *J. Biol. Chem.* 2004; 279:10593–10605. [PubMed: 14676205]
7. Salehi AH, Xanthoudakis S, Barker PA. *J. Biol. Chem.* 2002; 277:48043–48050. [PubMed: 12376548]
8. Salehi AH, et al. *Neuron.* 2000; 27:279–288. [PubMed: 10985348]
9. Hidaka K, et al. *J. Biol. Chem.* 2002; 277:32730–32738. [PubMed: 12105228]
10. Dubois E, et al. *J. Biol. Chem.* 2002; 277:23755–23763. [PubMed: 11956213]
11. Morrison BH, Bauer JA, Kalvakolanu DV, Lindner DJ. *J. Biol. Chem.* 2001; 276:24965–24970. [PubMed: 11337497]
12. Amundadottir LT, et al. *Nat. Genet.* 2006; 38:652–658. [PubMed: 16682969]
13. Haiman CA, et al. *Nat. Genet.* 2007; 39:638–644. [PubMed: 17401364]

Table 1

Association results for rs5945572 on Xp11.22 and prostate cancer in Iceland, The Netherlands, Spain, Sweden and the United States

Study population (cases/controls)	Frequency		OR (95% c.i.)	P value
	Cases	Controls		
Iceland (1,833/21,372)^a				
rs5945572 (A)	0.414	0.368	1.21 (1.09 – 1.34)	3.36 × 10 ^{−4}
Holland (991/2,021)				
rs5945572 (A)	0.390	0.347	1.20 (1.04 – 1.39)	0.015
Spain (539/1,594)				
rs5945572 (A)	0.432	0.364	1.33 (1.10 – 1.61)	3.20 × 10 ^{−3}
Sweden (2,865/1,722)				
rs5945572 (A)	0.421	0.379	1.19 (1.05 – 1.34)	4.84 × 10 ^{−3}
Baltimore, Maryland (1,516/554)				
rs5945572 (A)	0.397	0.334	1.31 (1.07 – 1.60)	8.39 × 10 ^{−3}
Chicago, Illinois (656/503)				
rs5945572 (A)	0.402	0.354	1.23 (0.98 – 1.54)	0.072
Nashville, Tennessee (526/612)				
rs5945572 (A)	0.392	0.353	1.18 (0.93 – 1.50)	0.18
Rochester, Minnesota (1,128/500)^b				
rs5945572 (A)	0.381	0.306	1.40 (1.11 – 1.76)	4.39 × 10 ^{−3}
All excluding Iceland (8,221/7,506)^c				
rs5945572 (A)	–	0.348	1.24 (1.16 – 1.33)	2.57 × 10 ^{−10}
All combined (10,054/28,879)^c				
rs5945572 (A)	–	0.351	1.23 (1.16 – 1.30)	3.95 × 10 ^{−13}

All *P* values shown are two-sided. Shown are the corresponding numbers of cases and controls (*n*), allelic frequencies of variants in affected and control individuals, the allelic odds ratio (OR) with 95% confidence interval (95% c.i.) and *P* values. For additional information on assays and frequencies, see Supplementary Tables 4 and 5 online.

^aResults presented for Iceland were adjusted for relatedness using a method of genomic control (see Supplementary Methods).

^bResults presented for the Rochester study group were adjusted for relatedness by applying a pedigree correction factor (see Supplementary Methods).

^cFor the combined study populations, the reported control frequency was the average, unweighted control frequency of the individual populations, whereas the OR and the *P* value were estimated using the Mantel-Haenszel model.

Table 2

Association results for rs721048 on 2p15 and prostate cancer in Iceland, The Netherlands, Spain, Sweden and the United States

Study population (cases/controls)	Frequency		OR (95% c.i.)	P value
	Cases	Controls		
Iceland(1,854/21,064)^a				
rs2710646 (A) ^b	0.227	0.202	1.16 (1.06 – 1.26)	7.79 × 10 ^{−4}
rs721048 (A)	0.229	0.204	1.16 (1.06 – 1.27)	9.22 × 10 ^{−4}
The Netherlands (998/2,021)				
rs721048 (A)	0.205	0.187	1.12 (0.98 – 1.28)	0.095
Spain (548/1,616)				
rs721048 (A)	0.234	0.210	1.15 (0.98 – 1.35)	0.088
Sweden (2,849/1,731)				
rs721048 (A)	0.201	0.186	1.10 (0.98 – 1.23)	0.090
Baltimore, Maryland (1,521/557)				
rs721048 (A)	0.205	0.206	1.00 (0.46 – 2.20)	0.99
Chicago, Illinois (665/552)				
rs721048 (A)	0.198	0.168	1.22 (1.00 – 1.49)	0.055
Nashville, Tennessee (526/612)				
rs721048 (A)	0.240	0.175	1.49 (1.21 – 1.83)	1.39 × 10 ^{−4}
Rochester, Minnesota (1,132/501)^c				
rs721048 (A)	0.224	0.184	1.28 (1.06 – 1.55)	0.012
All excluding Iceland (8,239/7,590)^d				
rs721048 (A)	–	0.188	1.15 (1.09 – 1.22)	2.23 × 10 ^{−6}
All combined (10,093/28,654)^d				
rs721048 (A)	–	0.190	1.15 (1.10 – 1.21)	7.66 × 10 ^{−9}

All *P* values shown are two-sided. Shown are the corresponding numbers of cases and controls (*n*), allelic frequencies of variants in affected and control individuals, the allelic odds ratio (OR) with 95% confidence interval (95% c.i.) and *P* values based on the multiplicative model. For additional information on assays and frequencies, see Supplementary Tables 4 and 5 online.

^aResults presented for Iceland were adjusted for relatedness using a method of genomic control (see Supplementary Methods).

^bThe SNPs rs2710646 and rs721048 are highly correlated ($r^2 = 0.99$), but rs2710646 failed in genotyping in some of the non-Icelandic groups, and results are therefore only presented for this marker from the Icelandic study group.

^cResults presented for the Rochester study group were adjusted for relatedness by applying a pedigree correction factor (see Supplementary Methods).

^dFor the combined study populations, the reported control frequency was the average, unweighted control frequency of the individual populations, whereas the OR and the *P* value were estimated using the Mantel-Haenszel model.