

A meta-analysis of genome-wide association studies to identify prostate cancer susceptibility loci associated with aggressive and non-aggressive disease

Ali Amin Al Olama^{1,†}, Zsofia Kote-Jarai^{2,†}, Fredrick R. Schumacher^{3,†}, Fredrik Wiklund^{4,†}, Sonja I. Berndt^{5,6,†}, Sara Benlloch^{1,‡}, Graham G. Giles^{7,8,‡}, Gianluca Severi^{7,8,‡}, David E. Neal^{9,10,‡}, Freddie C. Hamdy^{11,13,‡}, Jenny L. Donovan^{14,‡}, David J. Hunter^{15,‡}, Brian E. Henderson^{3,‡}, Michael J. Thun^{17,‡}, Michael Gaziano^{18,19,‡}, Edward L. Giovannucci^{16,‡}, Afshan Siddiq^{20,‡}, Ruth C. Travis^{12,‡}, David G. Cox^{21,22,‡}, Federico Canzian^{23,‡}, Elio Riboli^{21,‡}, Timothy J. Key^{12,‡}, Gerald Andriole^{26,‡}, Demetrius Albanes^{5,‡}, Richard B. Hayes^{27,‡}, Johanna Schleutker^{28,32,‡}, Anssi Auvinen^{30,‡}, Teuvo L.J. Tammela^{31,‡}, Maren Weischer^{33,‡}, Janet L. Stanford^{35,36,‡}, Elaine A. Ostrander^{37,‡}, Cezary Cybulski^{38,‡}, Jan Lubinski^{38,‡}, Stephen N. Thibodeau^{39,‡}, Daniel J. Schaid^{39,‡}, Karina D. Sorensen^{40,‡}, Jyotsna Batra^{42,‡}, Judith A. Clements^{42,‡}, Suzanne Chambers^{43,44,45,‡}, Joanne Aitken^{44,‡}, Robert A. Gardiner^{45,‡}, Christiane Maier^{46,47,‡}, Walther Vogel^{47,‡}, Thilo Dörk^{48,‡}, Hermann Brenner^{24,‡}, Tomonori Habuchi^{49,‡}, Sue Ingles^{3,‡}, Esther M. John^{50,51,‡}, Joanne L. Dickinson^{52,‡}, Lisa Cannon-Albright^{53,54,‡}, Manuel R. Teixeira^{55,‡}, Radka Kaneva^{56,‡}, Hong-Wei Zhang^{57,‡}, Yong-Jie Lu^{58,‡}, Jong Y. Park^{59,‡}, Kathleen A. Cooney^{60,61,‡}, Kenneth R. Muir^{62,‡}, Daniel A. Leongamornlert², Edward Saunders², Malgorzata Tymrakiewicz², Nadiya Mahmud², Michelle Guy², Koveela Govindasami², Lynne T. O'Brien², Rosemary A. Wilkinson², Amanda L. Hall², Emma J. Sawyer², Tokhir Dadaev², Jonathan Morrison¹, David P. Dearnaley^{2,63}, Alan Horwich^{2,63}, Robert A. Huddart^{2,63}, Vincent S. Khoo^{2,63}, Christopher C. Parker^{2,63}, Nicholas Van As⁶³, Christopher J. Woodhouse⁶³, Alan Thompson⁶³, Tim Dudderidge⁶³, Chris Ogden⁶³, Colin S. Cooper², Artitaya Lophatonanon⁶², Melissa C. Southey⁶⁴, John L. Hopper⁶⁵, Dallas English^{65,66}, Jarmo Virtamo⁶⁷, Loic Le Marchand⁶⁸, Daniele Campa²², Rudolf Kaaks²⁵, Sara Lindstrom¹⁵, W. Ryan Diver¹⁷, Susan Gapstur¹⁷, Meredith Yeager^{5,6}, Angela Cox⁶⁹, Mariana C. Stern³, Roman Corral³, Markus Aly^{4,70}, William Isaacs⁷², Jan Adolfsson⁷¹, Jianfeng Xu⁷³, S. Lilly Zheng⁷³, Tiina Wahlfors³¹, Kimmo Taari⁷⁴, Paula Kujala²⁹, Peter Klarskov³⁴, Børge G. Nordestgaard³³, M. Andreas Røder⁷⁵, Ruth Frikke-Schmidt⁷⁶, Stig E. Bojesen³³, Liesel M. FitzGerald³⁵, Suzanne Kolb³⁵, Erika M. Kwon³⁷, Danielle M. Karyadi³⁷, Torben Falck Orntoft⁴⁰, Michael Borre⁴¹, Antje Rinckleb⁴⁶, Manuel Luedeke⁴⁷, Kathleen Herkommer⁷⁷, Andreas Meyer⁴⁸, Jürgen Serth⁴⁸, James R. Marthick⁵², Briony Patterson⁵², Dominika Wokolorczyk³⁸, Amanda Spurdle⁷⁹, Felicity Lose⁷⁸, Shannon K. McDonnell³⁹, Amit D. Joshi³, Ahva Shahabi³, Pedro Pinto⁵⁵, Joana Santos⁵⁵, Ana Ray^{60,61}, Thomas A. Sellers⁵⁹, Hui-Yi Lin⁵⁹, Robert A. Stephenson⁷⁹, Craig Teerlink⁵³, Heiko Muller²⁴, Dietrich Rothenbacher²⁴,

[†]These authors jointly directed this work.

[‡]These authors contributed jointly to this work.

Norihiko Tsuchiya⁴⁹, Shintaro Narita⁴⁹, Guang-Wen Cao⁵⁷, Chavdar Slavov⁸⁰, Vanio Mitev⁵⁶, The UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology[†], The UK ProtecT Study Collaborators[§], The Australian Prostate Cancer Bioresource^{||}, The PRACTICAL Consortium[#], Stephen Chanock^{6,†}, Henrik Gronberg^{4,†}, Christopher A. Haiman^{3,†}, Peter Kraft^{15,†}, Douglas F. Easton^{1,†} and Rosalind A. Eeles^{2,63,*},[†]

¹Strangeways Laboratory, Centre for Cancer Genetic Epidemiology, Worts Causeway, Cambridge CB1 8RN, UK, ²The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK, ³Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Centre, Los Angeles, CA, USA, ⁴Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm SE-171 77, Sweden, ⁵Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD 20892, USA, ⁶Core Genotyping Facility, SAIC-Frederick, Inc., National Cancer Institute, NIH, Gaithersburg, MD, USA, ⁷Cancer Epidemiology Centre, Cancer Council Victoria, 1 Rathdowne Street, Carlton, VIC 3053, Australia, ⁸Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, 723 Swanston Street, Carlton, VIC 3053, Australia, ⁹Surgical Oncology (Uro-Oncology: S4), Addenbrooke's Hospital, University of Cambridge, Box 279, Hills Road, Cambridge, UK, ¹⁰Li Ka Shing Centre, Cancer Research UK Cambridge Research Institute, Cambridge CB2 2QQ, UK, ¹¹Nuffield Department of Surgery and ¹²Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK, ¹³Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK, ¹⁴School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK, ¹⁵Program in Molecular and Genetic Epidemiology, Department of Epidemiology and ¹⁶Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA, ¹⁷Epidemiology Research Program, American Cancer Society, Atlanta, GA 30303, USA, ¹⁸Massachusetts Veterans Epidemiology and Research Information Center (MAVERIC) and Geriatric Research, Education, and Clinical Center (GRECC), Boston Veterans Affairs Healthcare System, Boston, MA 02114, USA, ¹⁹Division of Aging, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA, ²⁰Department of Genomics of Common Disease, School of Public Health, Imperial College, London SW7 2AZ, UK, ²¹Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK, ²²Lyon Cancer Research Center, INSERM U1052, Lyon, France, ²³Genomic Epidemiology Group, ²⁴Division of Clinical Epidemiology and Aging Research and ²⁵Division of Cancer Epidemiology, German Cancer Research Centre (DKFZ), Heidelberg, Germany, ²⁶Division of Urologic Surgery, Washington University School of Medicine, St Louis, MO, USA, ²⁷Division of Epidemiology, Department of Environmental Medicine, NYU Langone Medical Centre, NYU Cancer Institute, New York, NY 10016, USA, ²⁸Institute of Biomedical Technology/BioMediTech, University of Tampere and ²⁹Department of Pathology, Centre for Laboratory Medicine, Tampere University Hospital, Tampere, Finland, ³⁰Department of Epidemiology, School of Health Sciences and ³¹Department of Urology, Tampere University Hospital and Medical School, University of Tampere, Tampere, Finland, ³²Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland, ³³Department of Clinical Biochemistry and ³⁴Department of Urology, Herlev Hospital, Copenhagen University Hospital, Herlev Ringvej 75, Herlev DK-2730, Denmark, ³⁵Division of Public Health Sciences, Fred Hutchinson Cancer Research Centre, Seattle, WA, USA, ³⁶Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA, ³⁷National Human Genome Research Institute, National Institutes of Health, 50 South Drive, Room 5351, Bethesda, MD, USA, ³⁸Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland, ³⁹Mayo Clinic, Rochester, MN, USA, ⁴⁰Department of Molecular Medicine and ⁴¹Department of Urology, Aarhus University Hospital, Skejby, Denmark, ⁴²Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation, Queensland University of Technology,

*To whom correspondence should be addressed at: Oncogenetics Team, The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK. Tel: +44 2086613642; Fax: +44 2087701489; Email: rosalind.eeles@icr.ac.uk

[†]UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology—membership list provided in Supplementary Material, Notes.

[§]UK ProtecT Study Collaborators—membership list provided in Supplementary Material, Notes.

^{||}Full list of participants of The Australian Prostate Cancer BioResource is provided in Supplementary Material, Notes.

[#]Full list of participants of The PRACTICAL Consortium is provided in the Supplementary Material, Notes.

[†]These authors jointly directed this work.

Brisbane, QLD, Australia, ⁴³Griffith Health Institute, Griffith University, Gold Coast, QLD, Australia, ⁴⁴Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Brisbane, QLD, Australia, ⁴⁵Centre for Clinical Research, University of Queensland, Brisbane, QLD, Australia, ⁴⁶Department of Urology and ⁴⁷Institute of Human Genetics, University Hospital Ulm, Ulm, Germany, ⁴⁸Hannover Medical School, Hannover, Germany, ⁴⁹Department of Urology, Akita University Graduate School of Medicine, 1-1-1 Hondo, Akita 010-8543, Japan, ⁵⁰Cancer Prevention Institute of California, Fremont, CA, USA, ⁵¹Division of Epidemiology, Department of Health Research and Policy and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA, ⁵²Menzies Research Institute Tasmania, University of Tasmania, Hobart, TAS, Australia, ⁵³Division of Genetic Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA, ⁵⁴George E. Wahlen Department of Veterans Affairs Medical Centre, Salt Lake City, UT, USA, ⁵⁵Department of Genetics, Portuguese Oncology Institute and Biomedical Sciences Institute (ICBAS), Porto University, Porto, Portugal, ⁵⁶Molecular Medicine Centre, Department of Medical Chemistry and Biochemistry, Medical University of Sofia, 2 Zdrave St, Sofia 1431, Bulgaria, ⁵⁷Department of Epidemiology, Second Military Medical University, Shanghai, China, ⁵⁸Centre for Molecular Oncology and Imaging, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London, UK, ⁵⁹Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Centre, 12902 Magnolia Drive, Tampa, FL, USA, ⁶⁰Department of Internal Medicine and ⁶¹Department of Urology, University of Michigan Medical School, Ann Arbor, MI, USA, ⁶²Warwick Medical School, University of Warwick, Coventry, UK, ⁶³Royal Marsden NHS Foundation Trust, Fulham and Sutton, London and Surrey, UK, ⁶⁴Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Grattan street, Parkville, VIC, Australia, ⁶⁵Cancer Epidemiology Centre, Cancer Council Victoria, 1 Rathdowne Street, Carlton, VIC, Australia, ⁶⁶Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, 723 Swanston Street, Carlton, VIC, Australia, ⁶⁷Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland, ⁶⁸Epidemiology Program, University of Hawaii Cancer Centre, Department of Medicine, John A. Burns School of Medicine, University of Hawaii, Honolulu, HI, USA, ⁶⁹Department of Oncology, University of Sheffield, Sheffield, UK, ⁷⁰Division of Urology, Department of Clinical Sciences, Danderyd Hospital and ⁷¹Oncological Centre, CLINTEC, Karolinska Institute, Stockholm, Sweden, ⁷²School of Medicine, Johns Hopkins University, 115 Marburg Building, 600 North Wolfe Street, Baltimore, MD 21205, USA, ⁷³Center for Cancer Genomics, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA, ⁷⁴Department of Urology, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland, ⁷⁵Department of Urology and ⁷⁶Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, Copenhagen DK-2100, Denmark, ⁷⁷Department of Urology, Rechts der Isar Medical Centre, Technical University of Munich, Munich, Germany, ⁷⁸Molecular Cancer Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, QLD, Australia, ⁷⁹Huntsman Cancer Institute, Salt Lake City, UT, USA and ⁸⁰Department of Urology and Alexandrovska University Hospital, Medical University of Sofia, Sofia, Bulgaria

Received June 20, 2012; Revised and Accepted October 4, 2012

Genome-wide association studies (GWAS) have identified multiple common genetic variants associated with an increased risk of prostate cancer (PrCa), but these explain less than one-third of the heritability. To identify further susceptibility alleles, we conducted a meta-analysis of four GWAS including 5953 cases of aggressive PrCa and 11 463 controls (men without PrCa). We computed association tests for approximately 2.6 million SNPs and followed up the most significant SNPs by genotyping 49 121 samples in 29 studies through the international PRACTICAL and BPC3 consortia. We not only confirmed the association of a PrCa susceptibility locus, rs11672691 on chromosome 19, but also showed an association with aggressive PrCa [odds ratio = 1.12 (95% confidence interval 1.03–1.21), $P = 1.4 \times 10^{-8}$]. This report describes a genetic variant which is associated with aggressive PrCa, which is a type of PrCa associated with a poorer prognosis.

INTRODUCTION

Genome-wide association studies (GWAS) have identified more than 50 common variants associated with susceptibility

to prostate cancer (PrCa). However, these variants explain less than a third of the familial risk of the disease, indicating that further susceptibility loci remain to be identified. Moreover, few variants identified by GWAS have thus far been

shown to be associated with aggressive PrCa. Although the current set of SNPs contributes to overall PrCa risk prediction, overall they do not discriminate men who will develop aggressive disease, a clinically more relevant outcome. A recent GWAS of aggressive PrCa identified a novel susceptibility locus on 2q37.3, but the per-allele odds ratio (OR) did not differ between aggressive and non-aggressive cases in the replication stage (1). It has been also reported that a genetic variant in *DAP2IP* might be associated with the risk of aggressive PrCa (2). A recent GWAS and validation study of aggressive PrCa found an SNP on 15q13, rs6497287, to be uniquely associated with this disease trait ($P_{\text{replication}} = 0.004$); however, lack of power due to small numbers and a non-significant test for heterogeneity between less and more aggressive PrCa warrants further investigation of this finding (3). Lin *et al.* (4) have reported an association of PrCa mortality with five germline SNPs from a *candidate* gene analysis.

RESULTS

In an attempt to identify susceptibility loci for aggressive PrCa, we conducted a meta-analysis of four GWAS (Table 1 and Supplementary Material, Notes). We also included data from the second stage of the UK study, that was genotyped for 43 671 SNPs showing evidence for association in stage 1 (5). These studies included, after quality control (QC) exclusions (see Materials and Methods), a total of 11 463 controls and 11 085 cases. For the present analysis, we included data from 5953 cases with aggressive disease defined as having a Gleason score of 8 or greater (with the exception of the BPC3 study, which also includes cases with tumor stage C or greater, and the CGEMS study, which also included cases with a Gleason score of 7) and all controls (it was ensured that there was no overlap between the studies). Following imputation using HapMap Phase II CEU as a reference, approximately 2.6 million genotyped and imputed SNPs were assessed in each GWAS study using a 1 df trend tests for association. Combined association tests were generated using a fixed effects meta-analysis (see Materials and Methods).

In the combined analysis, two loci, rs11672691 on 19q13 ($P\text{-value} = 3.8 \times 10^{-7}$) and rs11704416 on 22q13 ($P = 7.0 \times 10^{-6}$), showed strong evidence for association. rs11672691 is in the same region as rs887391 ($r^2 = 0.9$) that it was previously reported to be associated with PrCa by Hsu *et al.* (6), but it did not reach GWAS significance level in that report. These two SNPs were selected for further replication analysis in two international consortia, PRACTICAL and BPC3. The present analysis was restricted to 24 395 cases (2008 aggressive) and 24 726 controls (17 445 controls in aggressive disease analysis) from 26 studies from European populations (Table 1, Supplementary Material, Table S1 and Supplementary Material, Notes show all 29 studies, 26 of which are European).

SNP rs11672691 showed evidence of replication ($P = 0.006$) with a genome-wide significance of $P = 1.4 \times 10^{-8}$ in a combined analysis across all stages (Table 2) for aggressive PrCa. When data from non-aggressive cases were also included, the overall evidence for association was stronger ($P = 2.2 \times 10^{-12}$, overall). The per-allele OR for aggressive

Table 1. Total number of cases and controls included in the analysis

Study	Total number	Controls	Cases	Aggressive disease cases
GWAS meta-analysis				
Stage 1 UK	3748	1894	1854	617
Stage 2 UK/Melbourne	7590	3940	3650	1084
CGEMS	2277	1101	1176	688
CAPS	2926	994	1932	1091
BPC3	6007	3534	2473	2473
Total GWAS	22 548	11 463	11 085	5953
Confirmation				
PRACTICAL	34 188	17 324	16 864	1956
BPC3	14 933	7402	7531	52
Total confirmation	49 121	24 726	24 395	2008
Total all	71 669	36 189	35 480	7961

PrCa in the replication stage [1.12, 95% confidence interval (CI) 1.03–1.21; $P = 0.006$] was higher than that for non-aggressive cases (OR 1.08, 95% CI 1.05–1.12; $P = 8.2 \times 10^{-7}$); however, the difference was not statistically significant ($P\text{-value} = 0.18$). SNP rs11704416 showed evidence of replication for all PrCa ($P = 0.002$), but did not quite reach genome-wide significance overall ($P = 3.7 \times 10^{-7}$). The evidence for association with aggressive disease was weaker ($P = 0.16$ in the replication, $P = 4.0 \times 10^{-6}$ overall). There was no evidence that either locus was associated with serum PSA (based on 1578 control samples; Supplementary Material, Table S2). SNP rs11672691 showed stronger effect ($P = 0.02$) when we compared cases with a family history of PrCa (OR 1.14, 95% CI 1.06–1.22) with those with no family history (OR 1.06, 95% CI 1.02–1.10). The per-allele ORs did not differ significantly by ages (Supplementary Material, Table S3). Considering the estimated ORs in the replication stage, rs11672691 and rs11704416 together explain $\sim 0.16\%$ of the familial risk of PrCa.

DISCUSSION

rs11672691 lies between *ATP5SL* and *CEACAM21* (Fig. 1A) and within a hypothetical locus, *LOC100505495*, of a non-coding RNA. *ATP5SL* codes for an ATP synthase-like protein whose function is unknown; however, a variant in this gene has been associated with adult height (7). The carcinoembryonic antigen (*CEA*) gene family belongs to the immunoglobulin super family of genes. Several *CEA* subgroup members possess cell adhesion properties and some seem to function in signal transduction or regulation of signal transduction, possibly in association with other *CEA* sub-family members (8). Several of these proteins show a complex expression pattern in normal and cancerous tissues. Both *CEACAM5* and *CEACAM6* have a role in cell adhesion, invasion and metastasis (9), and are known to be overexpressed in a majority of carcinomas, including those of the gastrointestinal tract, the respiratory and genitourinary systems and breast cancer. The closest gene, *CEACAM21*, has been considered as a candidate gene for type 1 diabetes (10). A region on 19q13 (HPCQTL19) has been reported previously in a genetic linkage study to be a QTL for aggressive PrCa when the

Table 2. Meta-analysis and replication results

SNP chromosome allele position	Analysis	Study	OR ^a (95% CI)	P-value	P-value combined	P-value combined all
rs11672691 19 G/A 46677427	Aggressive disease cases	Stage 1 UK	1.20 (1.05–1.35)	0.02	3.8 × 10 ⁻⁷	1.4 × 10 ⁻⁸
		Stage 2 UK/Melbourne	NA ^b	NA ^b		
		CGEMS	1.25 (1.09–1.41)	0.006		
		CAPS	1.20 (1.05–1.35)	0.015		
		BPC3	1.14 (1.05–1.23)	0.01		
		PRACTICAL replication	1.11 (1.02–1.20)	0.006		
		BPC3 replication	1.45 (.85–2.48)	0.17		
		Replication all	1.12 (1.03–1.21)			
		Stage 1 UK	1.11 (1.00–1.24)	0.05		
	Stage 2 UK/Melbourne	NA ^b	NA ^b			
	CGEMS	1.20 (1.05–1.38)	0.009			
	CAPS	1.23 (1.10–1.37)	0.002			
	BPC3	1.14 (1.05–1.23)	0.006			
	PRACTICAL replication	1.08 (1.04–1.12)	2.6 × 10 ⁻⁵			
	BPC3 replication	1.10 (1.04–1.16)	0.002			
	Replication all	1.08 (1.05–1.12)				
	rs11704416 22 G/C 38766919	Aggressive disease cases	Stage 1 UK	0.85 (0.69–1.02)	0.056	3.3 × 10 ⁻⁶
Stage 2 UK/Melbourne			0.87 (0.75–0.999)	0.03		
CGEMS			0.75 (0.57–0.93)	0.002		
CAPS			0.77 (0.60–0.94)	0.003		
BPC3			0.94 (0.84–1.04)	0.21		
PRACTICAL replication			0.95 (0.87–1.04)	0.28		
BPC3 replication			0.62 (0.39–1.00)	0.05		
Replication all			0.94 (0.86–1.02)			
Stage 1 UK			0.90 (0.80–1.00)	0.058	7.0 × 10 ⁻⁶	
Stage 2 UK/Melbourne		0.91 (0.84–0.99)	0.03			
CGEMS		0.81 (0.69–0.94)	0.006			
CAPS		0.80 (0.65–0.96)	0.005			
BPC3		0.94 (0.84–1.04)	0.21			
PRACTICAL replication		0.94 (0.91–0.98)	0.003			
BPC3 replication		0.96 (0.91–1.03)	0.26			
Replication all		0.95 (0.92–0.98)				

Cases not classified as aggressive were those without the features defined in the text for aggressive disease.

^aPer-allele OR for the first allele.

^bNA: imputation quality was poor (0.227) and this result was excluded.

Gleason score was used as a quantitative measure of tumor aggressiveness (11).

SNP rs11704416 lies upstream of *TNRC6B* on chromosome 22 (Fig. 1B). The TNRC6 (trinucleotide repeat containing 6) family of proteins have been shown to stably associate with argonaute (AGO) proteins. AGO proteins, through their association with small RNAs, perform a critical function in the effector step of RNA interference. TNRC6B protein has a role in translational inhibition through its binding to AGOs (12).

These results illustrate the value of combining GWAS to confirm candidate loci where the genome-wide significance threshold was not obtained, and improve power identifying susceptibility loci associated with sub-classifications of diseases. The original report by Hsu *et al.* (6) implicating the 19q13 region failed to reach genome-wide significance, whereas our findings verify a significant association. Although some samples overlap between the Hsu *et al.* report and our study, we expanded the discovery phase by incorporating the Stage 1 UK and Stage 2 UK/Melbourne participants and including additional samples in the replication stage. The identification of loci involved in PrCa aggressiveness has been hampered by relatively small sample sizes. The locus reported here is associated with both aggressive and non-aggressive diseases,

and is therefore likely to be useful in determining those with clinically significant PrCa. Identification of such loci would aid the understanding of the biology of PrCa progression and targeted screening based on genetic risk profiling for aggressive disease.

MATERIALS AND METHODS

Samples

The four GWAS data sets have been described previously (1,5,13,14) (Table 1). Analyses were based on the data sets following standard QC procedures as previously described (5). The replication stage included 25 072 cases (2160 cases with a Gleason score of 8+) and 25 536 controls (18 255 in aggressive disease analysis) from 29 PrCa case-control studies (Supplementary Material, Table S1 and Notes). All studies were approved by the appropriate ethics committees.

Genotyping

In BPC3 and PRACTICAL, genotyping of samples from 13 studies was performed by the KASPar assay

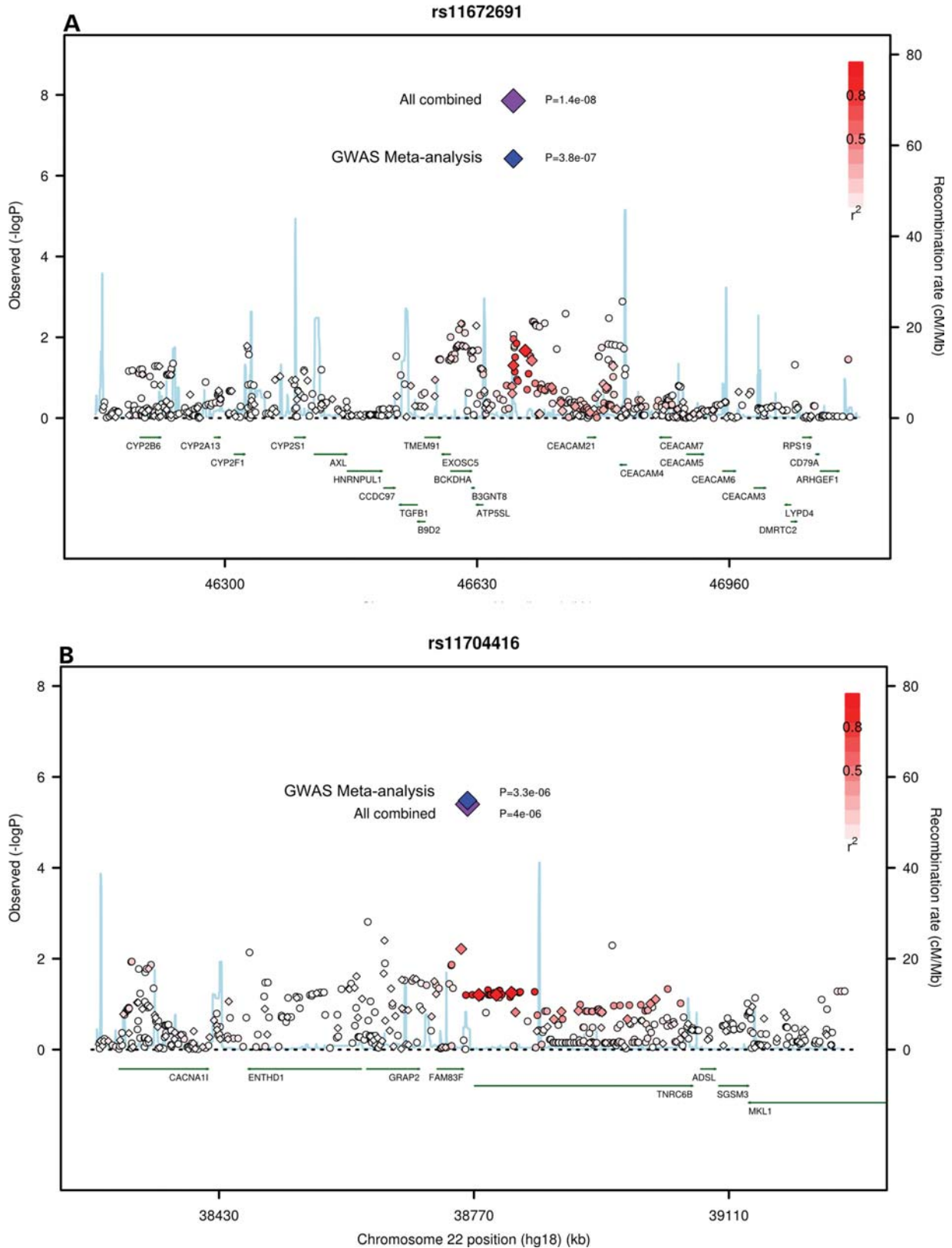


Figure 1. Regional plots of the associated SNPs at (A) 19q13 and (B) 22q13. Circles and diamonds denote imputed and genotyped SNPs, respectively. Plots show the genomic regions associated with PrCa and $-\log_{10}$ association P -values of SNPs. Also shown are the SNP build 36/hg18 coordinates in kilo bases, recombination rates and genes in the region. The intensity of red shading indicates the strength of LD (r^2) with the index SNP. Plots were drawn with a modified R script from <http://www.broadinstitute.org/mpg/snap/ldplot.php>.

(www.kbioscience.co.uk), whereas 15 study sites performed the 5' exonuclease assay (TaqMan™) using the ABI Prism 7900HT sequence detection system and IPO-Porto used TaqMan in a Roche LightCycler 480 Real-Time PCR System, all according to the manufacturer's instructions. Primers and probes were supplied directly by Applied Biosystems as Assays-By-Design™. Genotype counts are shown in Supplementary Material, Table S4. Assays at all sites included at least four negative controls and 2–5% duplicates on each 384-well plate. QC guidelines were followed by all the participating groups as described previously. In addition, all sites also genotyped 16 CEPH samples. We excluded individuals whose genotypes failed for at least 20% of the SNPs attempted. Data on a given SNP for a given site had to fulfill the following to be included: SNP call rate >95%, no deviation from Hardy–Weinberg equilibrium in controls at $P < 0.00001$, <2% discordance between genotypes in duplicate samples and in the CEPH samples. Cluster plots for SNPs that were close to failing any of the QC criteria were re-examined centrally.

Statistical methods

Imputation

Genotypes were imputed for approximately 2.6 million SNPs using the HapMap phase 2 CEU population as a reference. UK stages 1 and 2 and CGEMS were imputed using MACH 1.0 (<http://www.sph.umich.edu/csg/abecasis/MACH/>) to impute genotypes of autosomal markers and IMPUTE v1 (15) for chromosome X. The imputation for the BPC3 study was performed using MACH 1.0. The CAPS study used IMPUTE v1. We included imputed data from an SNP in the combined analysis if the estimated correlation between the genotype scores and the true genotypes (r^2) was >0.3 (MACH) or where the quality information was >0.3 (IMPUTE).

Analyses

For UK stages 1 and 2 and CGEMS, the imputed genotype probabilities were used to derive a 1 df association score statistic for each SNP, and its corresponding variance. The test statistic for UK stage 2 was stratified by population as described previously (5). In the BPC3 study, estimated betas and standard errors were calculated for each component study, including one principal component as a covariate to adjust for population structure, using ProbABEL (16), and the results were combined to generate overall betas and standard errors, using a fixed effects meta-analysis. CAPS used SNPTEST (https://mathgen.stats.ox.ac.uk/genetics_software/snpctest/snpctest.html) to estimate betas and standard errors. We converted the results from all studies into test scores and variances and derived a combined χ^2 trend statistic for each SNP (equivalent to the Mantel extension test, or as in a fixed effects meta-analysis) in the R package. We repeated the same procedure to combine the results for the case/control association analysis and aggressive case/control association analysis.

We assessed associations between each SNP and PrCa in the replication stage, using a 1 df Cochran–Armitage trend test stratified by study. The combined P -values over all stages were generated similarly (using a 1 df trend test based

on summing the scores and variances from each stage). SNPs were selected for validation on the basis of a significance level of $P < 10^{-7}$ in a combined meta-analysis of UK stages 1 and 2, CGEMS, CAPS and BPC3, excluding SNPs that were correlated with known susceptibility SNPs (SNP rs11704416 was included since it reached $P < 10^{-7}$ in an initial analysis). A total of 1921 subjects of non-European ancestry (Asian and African-American) were excluded from all analyses. Analyses were performed based on 2008 aggressive disease PrCa cases (out of 24 395 cases) and 17 445 controls (out of 24 726 controls). OR and 95% CI were estimated using unconditional logistic regression, stratified by study. In the text, we have reported the combined tests of association over all stages in European populations, but have emphasized the OR estimates from the replication stage to minimize the effect of 'winner's curse'. ORs were computed separately for subsets of cases defined by family history, grade and age. Modification of the ORs by family history and grade was assessed using a case-only analysis, using the dichotomous variable as the endpoint (family history Yes versus No/Grade GS < 8 versus GS \geq 8). Modification of the ORs by grade as a continuous covariate, and by age, was assessed using a case-only analysis, using polytomous regression with SNP genotype (scored 0, 1, 2) as the endpoint. The associations between SNP genotypes and PSA level were assessed using linear regression after log-transformation of PSA level to correct for skewness. We performed analyses for both all cases and only aggressive cases of PrCa. Analyses were performed in R principally using GenABEL (17), SNPTEST and ProbABEL (16) and Stata.

Publication of GWAS data

The U19, which provides funding for this work, plans to post summary data from this study onto a share point hosted by the NIH, by the end of 2012.

URLS

<http://www.srl.cam.ac.uk/consortia/practical/index.html>. (last accessed date 17 October, 2012)

<http://www.cgems.cancer.gov/>. (last accessed date 17 October, 2012)

https://mathgen.stats.ox.ac.uk/genetics_software/snpctest/snpctest.html. (last accessed date 17 October, 2012)

<http://www.sph.umich.edu/csg/abecasis/MACH/>. (last accessed date 17 October, 2012)

<http://www.broadinstitute.org/mpg/snap/ldplot.php>. (last accessed date 17 October, 2012)

<http://epi.grants.cancer.gov/BPC3/>. (last accessed date 17 October, 2012)

<http://ki.se/ki/jsp/polopoly.jsp?d=13809&a=29862&l=en>. (last accessed date 17 October, 2012)

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

Acknowledgements are detailed in Supplementary Material, Notes.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by Cancer Research UK grants (grant numbers: C5047/A7357, C1287/A10118, C5047/A3354, C5047/A10692, C16913/A6135) and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant (grant number: 1 U19 CA 148537-01 (the GAME-ON initiative)). Further support is detailed in Supplementary Material, Notes.

REFERENCES

- Schumacher, F.R., Berndt, S.I., Siddiq, A., Jacobs, K.B., Wang, Z., Lindstrom, S., Stevens, V.L., Chen, C., Mondul, A.M., Travis, R.C. *et al.* (2011) Genome-wide association study identifies new prostate cancer susceptibility loci. *Hum. Mol. Genet.*, **20**, 3867–3875.
- Duggan, D., Zheng, S.L., Knowlton, M., Benitez, D., Dimitrov, L., Wiklund, F., Robbins, C., Isaacs, S.D., Cheng, Y., Li, G. *et al.* (2007) Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. *J. Natl Cancer Inst.*, **99**, 1836–1844.
- FitzGerald, L.M., Kwon, E.M., Conomos, M.P., Kolb, S., Holt, S.K., Levine, D., Feng, Z., Ostrander, E.A. and Stanford, J.L. (2011) Genome-wide association study identifies a genetic variant associated with risk for more aggressive prostate cancer. *Cancer Epidemiol. Biomarkers Prev.*, **20**, 1196–1203.
- Lin, D.W., FitzGerald, L.M., Fu, R., Kwon, E.M., Zheng, S.L., Kolb, S., Wiklund, F., Stattin, P., Isaacs, W.B., Xu, J. *et al.* (2011) Genetic variants in the LEPR, CRY1, RNASEL, IL4, and ARVCF genes are prognostic markers of prostate cancer-specific mortality. *Cancer Epidemiol. Biomarkers Prev.*, **20**, 1928–1936.
- Eeles, R.A., Kote-Jarai, Z., Giles, G.G., Olama, A.A., Guy, M., Jugurnauth, S.K., Mulholland, S., Leongamornlert, D.A., Edwards, S.M., Morrison, J. *et al.* (2008) Multiple newly identified loci associated with prostate cancer susceptibility. *Nat. Genet.*, **40**, 316–321.
- Hsu, F.C., Sun, J., Wiklund, F., Isaacs, S.D., Wiley, K.E., Purcell, L.D., Gao, Z., Stattin, P., Zhu, Y., Kim, S.T. *et al.* (2009) A novel prostate cancer susceptibility locus at 19q13. *Cancer Res.*, **69**, 2720–2723.
- Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S. *et al.* (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, **467**, 832–838.
- Hammarstrom, S. (1999) The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Semin. Cancer Biol.*, **9**, 67–81.
- Blumenthal, R.D., Leon, E., Hansen, H.J. and Goldenberg, D.M. (2007) Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. *BMC Cancer*, **7**, 2.
- Howson, J.M., Walker, N.M., Smyth, D.J. and Todd, J.A. (2009) Analysis of 19 genes for association with type I diabetes in the Type I Diabetes Genetics Consortium families. *Genes Immun.*, **10** (Suppl. 1), S74–S84.
- Slager, S.L., Schaid, D.J., Cunningham, J.M., McDonnell, S.K., Marks, A.F., Peterson, B.J., Hebring, S.J., Anderson, S., French, A.J. and Thibodeau, S.N. (2003) Confirmation of linkage of prostate cancer aggressiveness with chromosome 19q. *Am. J. Hum. Genet.*, **72**, 759–762.
- Baillat, D. and Shiekhatter, R. (2009) Functional dissection of the human TNRC6 (GW182-related) family of proteins. *Mol. Cell. Biol.*, **29**, 4144–4155.
- Yeager, M., Chatterjee, N., Ciampa, J., Jacobs, K.B., Gonzalez-Bosquet, J., Hayes, R.B., Kraft, P., Wacholder, S., Orr, N., Berndt, S. *et al.* (2009) Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat. Genet.*, **41**, 1055–1057.
- Zheng, S.L., Liu, W., Wiklund, F., Dimitrov, L., Balter, K., Sun, J., Adami, H.O., Johansson, J.E., Chang, B., Loza, M. *et al.* (2006) A comprehensive association study for genes in inflammation pathway provides support for their roles in prostate cancer risk in the CAPS study. *Prostate*, **66**, 1556–1564.
- Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
- Aulchenko, Y.S., Struchalin, M.V. and van Duijn, C.M. (2010) ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics*, **11**, 134.
- Aulchenko, Y.S., Ripke, S., Isaacs, A. and van Duijn, C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, **23**, 1294–1296.