Germline variation at 8q24 and prostate cancer risk in men of European ancestry

Marco Matejcic, Edward J. Saunders et al. #

Chromosome 8q24 is a susceptibility locus for multiple cancers, including prostate cancer. Here we combine genetic data across the 8q24 susceptibility region from 71,535 prostate cancer cases and 52,935 controls of European ancestry to define the overall contribution of germline variation at 8q24 to prostate cancer risk. We identify 12 independent risk signals for prostate cancer (p < 4.28 × 10⁻¹⁵), including three risk variants that have yet to be reported. From a polygenic risk score (PRS) model, derived to assess the cumulative effect of risk variants at 8q24, men in the top 1% of the PRS have a 4-fold (95% CI = 3.62–4.40) greater risk compared to the population average. These 12 variants account for ~25% of what can be currently explained of the familial risk of prostate cancer by known genetic risk factors. These findings highlight the overwhelming contribution of germline variation at 8q24 on prostate cancer risk which has implications for population risk stratification.
Prostate cancer (PCa) is the most common cancer among men in the US, with 161,360 new cases and 26,730 related deaths estimated in 2017. Familial and epidemiological studies have provided evidence of substantial heritability of PCa, and ~170 common risk loci have been identified through genome-wide association studies (GWAS). The susceptibility region on chromosome 8q24 has been shown to be a major contributor to PCa risk, with multiple variants clustered in five linkage disequilibrium (LD) blocks spanning ~600 Mb that are independently associated with risk. Many of these association signals reported at 8q24 have been replicated across racial/ethnic populations, pointing to common shared functional variants within 8q24. However, rare ancestry-specific variants have also been detected, which confer larger relative risks of PCa (odds ratios [ORs] >2.0) than common risk variants in the region and signify allelic heterogeneity in the contribution of germline variation at 8q24 to PCa risk across populations.

In the current study, we perform a comprehensive investigation of genetic variation across the 1.4 Mb cancer susceptibility region at 8q24 (127.6–129.0 Mb) in relation to PCa risk. We combine genotyped and imputed data from two large GWAS consortia (PRACTICAL/ELLIPSE OncoArray and iCOGS) including >124,000 individuals of European ancestry to search for novel risk variants, as well as to determine the overall contribution of genetic variation at 8q24 to PCa heritability. Our findings underscore the sizable impact of genetic variation in the 8q24 region in explaining inter-individual differences in PCa risk, with potential clinical utility for genetic risk prediction.

Results
Marginal and conditional association analysis. Genotype data from the Illumina OncoArray and iCOGS array and imputation to 1000 Genomes Project (1KGP) were generated among 71,535 PCa cases and 52,935 controls of European ancestry from 86 case-control studies (see Methods). Of the 5600 genotyped and imputed variants at 8q24 (127.6–129.0 Mb) with minor allele frequency (MAF) > 0.1% retained for analysis (see Methods), we identified 12 variants with conditional p-values from the Wald test between 2.93 × 10⁻¹⁷ and 4.28 × 10⁻¹⁵ (Table 1). None of the other variants were statistically significant at p < 5 × 10⁻⁸ after adjustment for the 12 independent hits (Fig. 1). The 8q24 region is shown in Supplementary Fig. 1. Of these 12 stepwise signals, three had alleles with extreme risk allele frequencies (RAFs) that conveyed large effects (rs77541621, RAF = 2%, OR = 1.85, 95% CI = 1.76–1.94; rs183373024, RAF = 1%, OR = 2.67, 95% CI = 2.43–2.93; rs190257175, RAF = 99%, OR = 1.60, 95% CI = 1.42–1.80). The remaining variants had RAfs between 0.11 and 0.92 and conditional ORs that were more modest and ranged from 1.10 to 1.37 (Table 1). For 6 of the 12 variants, the allele found to be positively associated with PCa risk was the predominant allele (i.e., >50% in frequency). For two variants, rs78511380 and rs190257175, the marginal associations were not genome-wide significant and substantially weaker than those in the conditional model. For rs78511380, the marginal OR was slightly protective (OR = 0.97; p = 0.027), but reversed direction and was highly statistically significant when conditioning on the other 11 variants (OR = 1.19; p = 3.5 × 10⁻¹⁸; Table 1).

Haplotypic analysis. The haplotype analysis showed an additive effect of the 12 independent risk variants consistent with that predicted in the single variant test; co-occurrence of the 8q24 risk alleles on the same allele does not further increase the risk of PCa (Supplementary Table 1). The unique haplotype carrying the reference allele for rs190257175 (GCTTAT, 0.5% frequency) is also the sole haplotype associated with a reduced risk of PCa, suggesting that having the C allele confers a protective effect. The reference allele for rs78511380 (A, 8% frequency) occurs on a haplotype in block 2 together with the risk alleles for rs190257175, rs72758579 and rs5013678 (haplotype GTTTAA, 8%) which obscures the positive association with the T allele of rs78511380. Thus, the marginal protective effect associated with the risk allele for rs78511380 reflects an increased risk associated with the occurrence on a risk haplotype with other risk alleles (Supplementary Table 1).

Table 1 Marginal and conditional estimates for genetic markers at 8q24 independently associated with prostate cancer risk

<table>
<thead>
<tr>
<th>Variant IDa</th>
<th>Positionb</th>
<th>Allele</th>
<th>RAFc</th>
<th>LD cluster</th>
<th>Conditional associationf</th>
<th>Marginal association</th>
<th>OR (95%CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OR (95%CI)</td>
<td>p-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1941299</td>
<td>12791307</td>
<td>T/C</td>
<td>0.68</td>
<td>block 1</td>
<td>1.10 (1.08–1.12)</td>
<td>7.30 × 10⁻²⁵</td>
<td>1.09</td>
<td>1.07–21</td>
</tr>
<tr>
<td>rs1487240</td>
<td>128021752</td>
<td>A/G</td>
<td>0.74</td>
<td>block 1</td>
<td>1.20 (1.17–1.22)</td>
<td>2.77 × 10⁻⁶⁶</td>
<td>1.16</td>
<td>1.14–54</td>
</tr>
<tr>
<td>rs77541621</td>
<td>128077146</td>
<td>A/G</td>
<td>0.02</td>
<td>block 2</td>
<td>1.85 (1.76–1.94)</td>
<td>2.93 × 10⁻¹³⁷</td>
<td>1.83</td>
<td>1.74–137</td>
</tr>
<tr>
<td>rs190257775</td>
<td>128103466</td>
<td>T/C</td>
<td>0.99</td>
<td>block 2</td>
<td>1.60 (1.42–1.80)</td>
<td>4.28 × 10⁻¹⁵</td>
<td>1.36</td>
<td>1.22–1.53</td>
</tr>
<tr>
<td>rs72758579</td>
<td>128103969</td>
<td>T/C</td>
<td>0.18</td>
<td>block 2</td>
<td>1.31 (1.28–1.35)</td>
<td>1.26 × 10⁻⁸³</td>
<td>1.17</td>
<td>1.14–960</td>
</tr>
<tr>
<td>rs5013678</td>
<td>128103979</td>
<td>T/C</td>
<td>0.78</td>
<td>block 2</td>
<td>1.10 (1.08–1.13)</td>
<td>1.58 × 10⁻¹⁰⁹</td>
<td>1.20</td>
<td>1.17–990</td>
</tr>
<tr>
<td>rs183373024</td>
<td>128104117</td>
<td>G/A</td>
<td>0.01</td>
<td>block 2</td>
<td>2.67 (2.43–2.93)</td>
<td>4.89 × 10⁻⁹⁵</td>
<td>3.20</td>
<td>2.92–350</td>
</tr>
<tr>
<td>rs78511380</td>
<td>128104146</td>
<td>T/A</td>
<td>0.92</td>
<td>block 2</td>
<td>1.19 (1.14–1.23)</td>
<td>3.48 × 10⁻¹⁸</td>
<td>0.97</td>
<td>0.94–100</td>
</tr>
<tr>
<td>rs17464492</td>
<td>128342866</td>
<td>A/G</td>
<td>0.72</td>
<td>block 2</td>
<td>1.16 (1.14–1.18)</td>
<td>3.01 × 10⁻⁵⁰</td>
<td>1.17</td>
<td>1.15–990</td>
</tr>
<tr>
<td>rs6982267</td>
<td>128413025</td>
<td>G/T</td>
<td>0.51</td>
<td>block 5</td>
<td>1.18 (1.16–1.20)</td>
<td>5.68 × 10⁻⁸⁴</td>
<td>1.23</td>
<td>1.21–315</td>
</tr>
<tr>
<td>rs7812894</td>
<td>128520479</td>
<td>A/T</td>
<td>0.13</td>
<td>block 5</td>
<td>1.37 (1.33–1.40)</td>
<td>1.55 × 10⁻¹²</td>
<td>1.45</td>
<td>1.41–149</td>
</tr>
<tr>
<td>rs12549761</td>
<td>128540776</td>
<td>C/G</td>
<td>0.87</td>
<td>block 5</td>
<td>1.21 (1.18–1.24)</td>
<td>1.61 × 10⁻⁴⁵</td>
<td>1.28</td>
<td>1.25–131</td>
</tr>
</tbody>
</table>

aVariants that remained genome-wide significantly associated with PCa risk (p < 5 × 10⁻⁸) in the final stepwise model
bChromosome position based on human genome build 37
cMinor allele frequency
dRisk allele/reference allele
eLD clusters were inferred based on recombination hotspots using Haploview 4.2 and defined as previously reported by Al Olama et al.
fEach variant was incorporated in the stepwise model based on the strength of marginal association from the meta-analysis of OncoArray and iCOGS data
ɡPer-allele odds ratio and 95% confidence interval adjusted for country, 7(OncoArray)/8(iCOGS) principal components and all other variants in the table
hPer-allele odds ratio and 95% confidence interval adjusted for country and 7(OncoArray)/8(iCOGS) principal components
Correlation with known risk loci. The 12 risk variants spanned across the five LD blocks previously reported to harbor risk variants for PCa at 8q24, with block 2 harboring six signals, blocks 1 and 5 two signals each, and blocks 3 and 4 only one (Supplementary Fig. 2). Except for a weak correlation between rs72725879 and rs78511380 in block 2 ($r^2 = 0.28$), the risk variants were uncorrelated with each other ($r^2 \leq 0.09$; Supplementary Data 1), which corroborates their independent association with PCa risk. Eight of the variants (rs1487240, rs77541621, rs72725879, rs5013678, rs183373024, rs17464492, rs6983267, rs7812894)
rs7812894) have been previously reported either directly (Supplementary Table 2) or are correlated ($r^2 \geq 0.42$) with known markers of PCa risk from studies in populations of European, African or Asian ancestry (Supplementary Data 1). The marginal estimates for previously published PCa risk variants at 8q24 in the current study are shown in Supplementary Table 2. The variant rs1914295 in block 1 is only weakly correlated at $r^2 = 0.17$ and 0.14, respectively, while rs7851380 is modestly correlated with the previously reported risk variant rs1016343 ($r^2 = 0.28$). The remaining two variants, rs190257175 and rs12549761, are not correlated ($r^2 < 0.027$) with any known PCa risk marker.

Polycigenic risk score and familial relative risk. To estimate the cumulative effect of germline variation at 8q24 on PCa risk, a polygenic risk score (PRS) was calculated for the 12 independent risk alleles from the final model based on allele dosages weighted by the per-allele conditionally adjusted ORs (see Methods). Compared to the men at ‘average risk’ (i.e., the 25th–75th PRS range among controls), men in the top 10% of the PRS distribution had a 1.93-fold relative risk (95% CI = 1.86–2.01) (Table 2), with the risk being 3.99-fold higher (95% CI = 3.62–4.40) for men in the top 1%. Risk estimates by PRS category were not modified by family history (FamHist-yes: OR = 4.24, 95% CI = 2.85–6.31; FamHist-no: OR = 3.38, 95% CI = 2.88–3.97). To quantify the impact of germline variation at 8q24, we also estimated the proportion of familial relative risk (FRR) and heritability of PCa contributed by 8q24 and compared this to the proportions explained by all known PCa risk variants including 8q24 (see Methods). The 175 established PCa susceptibility loci identified to date are estimated to explain 25.4% of the total FRR and 25.4% of the total heritability of PCa (Table 3). This is similar to the proportion of heritability explained by 8q24 variants (22.2%) compared to the total explained heritability by the known risk variants (0.118). In comparison, the next highest contribution of an individual susceptibility region to the FRR of PCa is the TERT region at chromosome 5p15, where 5 independent signals contributed 2.63% (95% CI = 2.34–3.00). No other individual GWAS...
locus has been established as explaining >2% of the FRR, including the low frequency, non-synonymous, moderate penetrance HOXB13 variant (rs138213197) at chromosome 17q21 that is estimated to explain only 1.91% (95%CI = 1.20–2.85) of the FRR.

JAM analysis. We explored our data with a second fine-mapping approach, JAM (Joint Analysis of Marginal summary statistics)\textsuperscript{12}, which uses GWAS summary statistics to identify credible sets of variants that define the independent association signals in susceptibility regions (see Methods). The 95% credible set for the JAM analysis confirmed all of the independent signals from stepwise analysis except rs190257175, for which evidence for an association was weak (variant-specific Bayes factor (BF) = 1.17). There were 50 total variants included in the 95% credible set, and 174 after including variants in high LD ($r^2 > 0.9$) with those in the credible set (Supplementary Data 2).

Discussion
In this large study of germline genetic variation across the 8q24 region, we identified 12 independent association signals among men of European ancestry, with three of the risk variants (rs1914295, rs190257175, and rs12549761) being weakly correlated ($r^2 \leq 0.17$) with known PCa risk markers. The combination of these 12 independent signals at 8q24 capture approximately one quarter of the total FRR explained by known genetic risk factors, which is substantially greater than any other known PCa risk locus.

The 8q24 region is the major susceptibility region for PCa; however, the underlying biological mechanism(s) through which germline variation in this region influences PCa risk remains uncertain. For each of the 12 risk variants at 8q24, the 95% credible set defined noteworthy (i.e., putative functional) variants based on summary statistics while accounting for LD. To inform biological functionality, we overlaid epigenetic functional annotation using publicly available datasets (see Methods) with the location of the 12 independent signals (and corresponding 174 variants within their 95% credible sets; Supplementary Data 3). Of the 12 independent lead variants, 6 are situated within putative transcriptional enhancers in prostate cell-lines; either through intersection with H3K27AC (rs72725879, rs5013678, rs78511380, rs6983267 and rs7812894) or through a ChromHMM enhancer annotation (rs17464492, rs6983267, rs7812894). Eight of the 12 stepwise hits (rs77541621, rs190257175, rs5013678, rs183373024, rs78511380, rs6983267, and rs7812894) also intersect transcription factor binding site peaks from multiple ChIP-seq datasets representing the AR, ERG, FOXA1, GABPA, GATA2, HOX13, and NKX3.1 transcription factors, with all 8 intersecting a FOXA1 mark and half an AR binding site. These variants may therefore exert their effect through regulation of enhancer activity and long-range expression of the proto-oncogene MYC in vitro and in vivo\textsuperscript{14,15}. However, despite the close proximity to the MYC locus, no direct association has been detected between 8q24 risk alleles and MYC expression in normal and tumor human prostate tissues\textsuperscript{16}. The rare variant with the largest effect on risk, rs183373024, shows high evidence of functionality based on overlap with multiple DNaseI and transcription factor binding site peaks (for AR, FOXA1, HOX13, and NKX3.1), which supports previous findings of an allele-dependent effect of this variant on the disruption of a FOXA1 binding motif\textsuperscript{17}. Seven independent signals (rs1914295, rs1487240, rs77541621, rs72725879, rs5013678, rs183373024, rs78511380) and variants correlated at $r^2 > 0.9$ with these signals (Supplementary Data 2) are located within or near a number of prostate cancer–associated long noncoding RNAs (lncRNAs), including PRNCR1, PCAT1, and CCAT2, previously reported to be upregulated in human PCa cells\textsuperscript{18} and tissues\textsuperscript{19,20}. Based on eQTL annotations in prostate adenocarcinoma cells, the independent signal rs1914295 and three correlated variants ($r^2 > 0.9$; Supplementary Data 2) are associated with overexpression of FAM54B, a gene previously associated with progression and poor prognosis of PCa in animal studies\textsuperscript{21}. Variants correlated at $r^2 > 0.9$ with rs7812894 ($n = 9$; Supplemental Table 4) are eQTLs for POUSF1B, a gene overexpressed in cancer cell lines and cancer tissues\textsuperscript{22,23}, although its role in PCa development is unknown. Whilst we have successfully refined the 8q24 region and identified a subset of variants with putative biological function within our credible set, multi-ethnic comparisons may help refine the association signals even further and precisely identify the functional alleles and biological mechanisms that modify PCa risk.

Whereas the individual associations of the 8q24 variants with PCa risk, at relative modest ORs ($< 2.0$, except for rs183373024), their cumulative effects are substantial, with risk being 4-fold higher for men in the top 1% of the 8q24-only PRS. The contribution to the overall FRR of PCa is substantially greater for the 8q24 region (9.42%) than for any other known GWAS locus, including the moderate penetrance non-synonymous variant in HOXB13 (1.91%). The ability of these markers to explain ~25.4% of what can be currently explained by all known PCa risk variants is a clear indication of the important contribution of germline variation at 8q24 on PCa risk. Our study was predominantly powered to analyze variants with MAF > 1% as the imputed variants with MAF = 0.1-1% were most likely to fail quality control (QC); however, the high density of genotyped markers and haplotypes at 8q24 in the OncoArray and iCOGS studies provided a robust backbone for imputation and increased the chances to impute lower MAF variants with high imputation quality score. Understanding of the biology of these variants and the underlying genetic basis of PCa could provide new insights into the identification of reliable risk-prediction biomarkers for PCa, as well as enable the development of effective strategies for targeted screening and prevention.

Methods

Study subjects, genotyping, and quality control. We combined genotype data from the PRACTICAL/ELLIPSE OncoArray and iCOGS consortia\textsuperscript{24,25}, which included 143,699 men of European ancestry from 86 case-control studies largely based in either the US or Europe. In each study, cases primarily included men with incident PCa while controls were men without a prior diagnosis of the disease.

Both of the OncoArray and iCOGS custom arrays were designed to provide high coverage of common alleles (minor allele frequency [MAF] > 5%) across 8q24 (127.6–129.0 Mb) based on the 1000 Genomes Project (1KGP) Phase 3 for OncoArray, and the European ancestry (EUR) panel from HapMap Phase 2 for iCOGS. A total of 57,580 PCa cases and 37,927 controls of European ancestry were genotyped with the Illumina OncoArray, and 24,198 PCa cases and 23,994 controls of European ancestry were genotyped with the Illumina iCOGS array. For both studies, sample exclusion criteria included duplicate samples, first-degree relatives, samples with a call rate <95% or with extreme heterozygosity ($p < 10^{-6}$), and samples with an estimated proportion of European ancestry <0.8\textsuperscript{34}. In total, genotype data for 53,449 PCa cases and 36,224 controls from OncoArray and 18,086 PCa cases and 16,711 controls from iCOGS were included in the analysis. Genetic variants with call rates <0.95, deviation from Hardy-Weinberg equilibrium ($p < 10^{-6}$ in controls), and genotype discrepancy in >2% of duplicate samples were excluded. Of the final 488,417 genotyped variants on the OncoArray and 201,598 on the iCOGS array that passed QC, 1581 and 1737 within the 8q24 region, respectively, were retained for imputation. All studies complied with all relevant ethical regulations and were approved by the institutional review boards at each of the participating institutions. Informed consent was obtained from all study participants. Additional details of each study are provided in the Supplementary Note 1.

Imputation analysis. Imputation of both OncoArray and iCOGS genotype data was performed using SHAPEIT\textsuperscript{25} and IMPUTeV\textsuperscript{26} to the October 2014 (Phase 3)
release of the 1KGP reference panel. A total of 10,136 variants from OncoArray and 10,360 variants from iCOGS with MAF > 0.1% were imputed across the risk region at 8q24 (127.6-129.0 Mb). Variants with an imputation quality score >0.8 were retained for a total of 5600 overlapping variants between the two datasets.

**Statistical analysis.** Unconditional logistic regression was estimated to test per-allele odds ratios (ORs) and 95% confidence intervals (CIs) for the association between genetic variants (single nucleotide polymorphisms and insertion/deletion polymorphisms (InDels)) and PCa risk adjusting for country and principal components (7 for OncoArray and 8 for iCOGS). Allele dosage effects were tested through a 1-degree of freedom two-tailed Wald trend test. The marginal risk estimates for the 5600 variants at 8q24 that passed QC were combined by a fixed effect meta-analysis with inverse variance weighting using METAL. A modified forward and backward stepwise model selection with inclusion and exclusion criteria (p < 0.05 and p > 0.10) was performed to identify variants marginally associated with PCa risk from the meta results (p < 0.05, n = 2772). At each step, the effect estimates for the candidate variants from both studies (OncoArray and iCOGS) were meta-analyzed and each variant was incorporated into the model based on the strength of association. All remaining variants were included one at a time into the logistic regression model conditioning on those already incorporated in the model. We applied a conservative threshold for independent associations, with variants kept in the model if their meta p-value from the Wald test was genome-wide significant at p ≤ 5 × 10−8 after adjustment for the other variants in the model. Correlations between genetic variants (single nucleotide polymorphisms and insertion/deletion polymorphisms (InDels)) in a mixture region at 8q24 (127.6-129.0 Mb). Variants with an imputation quality score >0.8 and 10,360 variants from iCOGS with MAF > 0.1% were imputed across the risk release of the 1KGP reference panel. A total of 10,136 variants from OncoArray and 8,782 from the meta-analysis results using LD estimated from imputed individual level datasets as putative evidence of biological functionality using publicly available datasets as described in Hoffman et al. Brieﬂy, variants were annotated for proximity to gene (GENCODEv19), miRNA transcripts (mirBase release 20), evolutionary constraint (according to GERP++, Siph and PhastCons algorithms), likelihood of pathogenicity (CADDv1.3) and overlap with prospective regulatory elements in prostate-speciﬁc datasets (DNase hypersensitivity sites, H3K27Ac, H3K27me3 and H3K4me3 histone modiﬁcations, and for AR, CTCF, ERG, FOXA1, GATA2, HOXB13, and NKX3.1 transcription factor binding sites) in a mixture of LNCaP, PC-3, PrEC, RWPE1, and VCaP cell lines and human prostate tumor tissues downloaded from the Cistrome Data Browser (http://cistrome.org/db/). The chromatin state in which each variant residues was assessed using ChromHMM annotations from two prostate cell lines (PrEC and PC3). Cis-gene regulation was evaluated using 359 prostate adenoma cases from The Cancer Genome Atlas (TCGA PRAD; https://gdc-portal.nci.nih.gov) that passed QC. The eQTL analysis was performed using FastQTL with 1000 permutations for each gene window. We then used the method by Nica et al. that integrates eQTLs and GWAS results in order to reveal the subset of association signals that are due to cis-eQTLs. For each significant eQTL, we added the candidate variant to the linear regression model to assess if the inclusion better explains the change in expression of the gene. We retrieved the p-value of the model, assigning p-value of 1 if the eQTL and variant are the same. Then we ranked the p-values in descending order for each eQTL and finally calculated the colocalization score for each pair of eQTLs and variants. In general, if an eQTL and candidate represent the same signal, this will be refl ected by the variant having a high p-value, a low rank and consequently a high colocalization score.

**Data availability**

The authors declare that data supporting the findings of this study are available in the manuscript and the supplementary information file. However, some of the data used to generate the results of this study are available from the first author and the FRAC-TICAL Consortium upon request.

Received: 7 February 2018 Accepted: 1 October 2018
Published online: 05 November 2018

**References**

Acknowledgements

Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) [U19 CA 148537 for ELAcidating Loci Involved in Prostate Cancer Susceptibility (ELLIPSE) project and R01HG07492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN2682012000081]. Additional analytic support was provided by NIH NC1 U01 CA188392 (Ph: Schumacher). The PRACTICAL consortium (http://practicalacr.ac.uk/) was supported by Cancer Research UK Grants C5047A73575, C5047A16536, C5047A116536, C5047A73534, C5047A10692, C16913/A6135, European Commission’s Seventh Framework Programme grant agreement n° 232175 (HEALTH-F2-2009-223175), and The National Institute of Health (NIH) Cancer Post–Cancer GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative). We wish to thank all GWAS study groups contributing to the data set from which this study was conducted: OncoArray, iCOGS, The PRACTICAL (Prostate Cancer Association) Group to Investigate Cancer–Associated Alterations in the Genome) Consortium; and The GAME-ON/ELLIPSE Consortium. Detailed acknowledgements and funding information for all GWAS study groups and from all the individual studies involved in the PRACTICAL Consortium are included in Supplementary Note 1. We would also like to thank the following for funding support: The Institute of Cancer Research and The Everyday Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now Prostate Action), The Orchid Cancer Appeal, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIH funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust.

Author contributions


Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06863-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
The PRACTICAL (Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome) Consortium


83Molecular Cancer Epidemiology Laboratory, QIMR Berghofer Institute of Medical Research, Herston, QLD 4006, Australia. 84School of Medicine, University of Queensland, Herston, QLD 4006, Australia. 85Royal Brisbane and Women’s Hospital, Herston, QLD 4029, Australia. 86The Kinghorn Cancer Centre (TKCC), Victoria, NSW 2010, Australia. 87Prostate Cancer Research Group, South Australian Health and Medical Research Institute, Adelaide, SA 5000, Australia. 88Department of Physiology, Biomedicine, and Molecular Biology, Umeå University, SE-901 85 Umeå, Sweden. 89Department of Urology, Helsinki University Central Hospital and University of Helsinki, FI-00014 Helsinki, Finland. 90Division of Biostatistics and Bioinformatics, University of Maryland Greenbaum Cancer Center, and Department of Public Health, University of Maryland School of Medicine, Baltimore, MD 21201, USA. 91Cancer Genomics Research Laboratory (CGR), Division of Cancer Epidemiology and Genetics, FNLCR Leidos Biomedical Research, National Cancer Institute, Frederick, MD 21701, USA. 92DNA Extraction and Staging Laboratory (DESL), Cancer Genomics Research Laboratory (CGR), Division of Cancer Epidemiology and Genetics, FNLCR Leidos Biomedical Research, National Cancer Institute, Frederick, MD 21701, USA. 93Sheffield Institute for Nucleic Acids, University of Sheffield, Sheffield S10 2TN, UK. 94Cambridge Cancer Trials Centre, Cambridge Clinical Trials Unit-Cancer Theme, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK. 95Department of Medical Imaging, University Health Network, Toronto, ON M5G 2C4, Canada. 96Department of Pathology, University Health Network, Toronto, ON M5G 2C4, Canada. 97Advanced Radiation Biology Research Program, Research Center for Charged Particle Therapy, National Institute of Radiological Sciences, Chiba 263-8555, Japan. 98Department of Urology, National Hospital Organization Tokyo Medical Center, Tokyo 152-8902, Japan. 99Department of Urology, Nanfang Hospital, Southern Medical University, 510515 Guangzhou, China. 100Department of Pathology, The First Affiliated Hospital, Zhejiang University Medical College, 310009 Hangzhou, China. 101Department of Pathology, Changhai Hospital, The Second Military Medical University, 200433 Shanghai, China. 102Department of Urology, First Affiliated Hospital, Medical College, Zhengzhou University, 450003 Zhengzhou, China. 103Department of Urology, North Sichuan Medical College, 637000 Nanchong, China. 104Department of Nutrition Science, Shenyang Medical College, 110034 Shenyang, China. 105Department of Urology, First Affiliated Hospital, Complexo Hospitalario Universitario de Santiago, SERGAS, 15706 Santiago de Compostela, Spain. 106Department of Urology, Hospital Germans Trias i Pujol, 08916 Barcelona, Spain. 107Laboratory and Department of Urology, Hospital Clinic, Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, 08036 Barcelona, Spain. 108Centre de Recerca Biomèdica CELLEX, 08036 Barcelona, Spain. 109Department and Laboratory of Urology, Hospital Clinic, Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, 08036 Barcelona, Spain. 110Department of Urology, Klinikum rechts der Isar der Technischen Universität München, 81675 Munich, Germany. 111Department of General and Clinical Pathology, Alexandrovskaya University Hospital, Medical University, 1431 Sofia, Bulgaria. 112Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah 21522 3270, Saudi Arabia. 113Grupo de Medicina Xénómica, CIBERER, CIMUS, Universidad de Santiago de Compostela, Avenida de Barcelona, 15782 Santiago de Compostela, Spain. 114Hérault Cancer Registry, Montpellier cedex 5, 34298 Montpellier, France. 115INserm U1147, 75013 Paris, France. 116Department of Clinical Science, Intervention and Technology, Karolinska Institutet, SE-171 77 Stockholm, Sweden. 117Swedish Agency for Health Technology Assessment and Assessment of Social Services, SE-102 33 Stockholm, Sweden. 118Department of Surgical and Perioperative Sciences, Urology and Andrology, Umeå University, SE-901 85 Umeå, Sweden. 119Department of Urology, Uppsala University, SE-751 85 Uppsala, Sweden. 120Department of Urology, Faculty of Medicine and Health, Örebro University, SE-701 82 Örebro, Sweden. 121Medical Research Council (MRC) Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK. 122Department of Health Research (NIHR) Biomedical Research Centre, University of Bristol, Bristol BS8 1TH, UK. 123Department of Urology, Cancer Therapy and Research Center, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA. 124Menzies Health Institute Queensland, Griffith University, Gold Coast, QLD 4222, Australia. 125Cancer Council Queensland, Fortitude Valley, QLD 4006, Australia. 126Chris O’Brien Lifehouse (COBLH), Camperdown, Sydney, NSW 2010, Australia. 127Garvan Institute of Medical Research, Sydney, NSW 2010, Australia. 128Dame Roma Mitchell Cancer Research Centre, University of Adelaide, Adelaide, SA 5005, Australia. 129Department of Anatomy and Developmental Biology, Biomedical Discovery Institute, Monash University, Melbourne, VIC 3800, Australia. 130Prostate Cancer Translational Research Program, Cancer Research Division, Peter MacCallum Cancer Centre, Melbourne, VIC 3000, Australia. 131Department of Molecular Medicine and Surgery, Karolinska Institutet, and Department of Urology, Karolinska University Hospital, 171 76 Stockholm, Sweden. 132Department of Clinical Sciences at Danderyd Hospital, Karolinska Institutet, 182 88 Stockholm, Sweden. 133Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. 134Department of Urology, Tampere University Hospital, University of Tampere, KALEVANTTI 4, FI-33014 Tampere, Finland. 135Faculty of Medicine and Life Sciences, University of Tampere, FI-33014 Tampere, Finland. 136Department of Epidemiology, School of Health Sciences, University of Tampere, FI-33014 Tampere, Finland. 137Cancer Council Queensland, Fortitude Valley, QLD 4006, Australia. 138Department of Urology, Aarhus University Hospital, 8200 Aarhus N, Denmark. 139Department of Radiation Oncology, University of Rochester Medical Center, Rochester, NY 14620, USA. 140Department of Pathology and Pediatrics, Albert Einstein College of Medicine, Bronx, NY 10461, USA. 141Second Military Medical University, 200433 Shanghai, China. 142Wuxi Second Hospital, Nanjing Medical University, 214003 Wuxi, Jiangsu, China. 143Department of Urology, The First Affiliated Hospital, Chongqing Medical University, 200032