Loveday, C; Litchfield, K; Levy, M; Holroyd, A; Broderick, P; Kote-Jarai, Z; Dunning, AM; Muir, K; Peto, J; Eeles, R; Easton, DF; Dudakia, D; Orr, N; Pashayan, N; Reid, A; Huddart, RA; Houlston, RS; Turnbull, C (2017) Validation of loci at 2q14.2 and 15q21.3 as risk factors for testicular cancer. Oncotarget, 9 (16). pp. 12630-12638. ISSN 1949-2553 DOI: https://doi.org/10.18632/oncotarget.23117

Downloaded from: http://researchonline.lshtm.ac.uk/4647132/

DOI: 10.18632/oncotarget.23117

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
Validation of loci at 2q14.2 and 15q21.3 as risk factors for testicular cancer

Chey Loveday, Kevin Litchfield, Max Levy, Amy Holroyd, Peter Broderick, Zsofia Kote-Jarai, Alison M Dunning, Kenneth Muir, Julian Peto, Rosalind Eeles, Alison M Dunning, Kenneth Muir, Julian Peto, Rosalind Eeles, Douglas F Easton, Darshna Dudakia, Nick Orr, Nora Pashayan, Alison Reid, Robert A Huddart, Richard Houlston and Clare Turnbull

1Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK
2Translational Cancer Therapeutics Laboratory, The Francis Crick Institute, London, UK
3Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK
4Division of Health Sciences, Warwick Medical School, Warwick University, Warwick, UK
5Institute of Population Health, University of Manchester, Manchester, UK
6Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
7Royal Marsden NHS Foundation Trust, London, UK
8Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
9The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK
10Department of Applied Health Research, University College London, London, UK
11Academic Uro-oncology Unit, The Royal Marsden NHS Foundation Trust, Sutton, Surrey, UK
12Academic Radiotherapy Unit, Institute of Cancer Research, Sutton, Surrey, UK
13William Harvey Research Institute, Queen Mary University, London, UK
14Guys and St Thomas NHS Foundation Trust, London, UK

Correspondence to: Clare Turnbull, email: clare.turnbull@icr.ac.uk

Keywords: testicular cancer; germ cell tumor; TGCT; GWAS; oncoarray

Received: July 12, 2017  Accepted: November 15, 2017  Published: December 07, 2017

ABSTRACT

Testicular germ cell tumor (TGCT), the most common cancer in men aged 18 to 45 years, has a strong heritable basis. Genome-wide association studies (GWAS) have proposed single nucleotide polymorphisms (SNPs) at a number of loci influencing TGCT risk. To further evaluate the association of recently proposed risk SNPs with TGCT at 2q14.2, 3q26.2, 7q36.3, 10q26.13 and 15q21.3, we analyzed genotype data on 3,206 cases and 7,422 controls. Our analysis provides independent replication of the associations for risk SNPs at 2q14.2 (rs2713206 at \( P = 3.03 \times 10^{-2} \); \( P \)-meta = 3.92 \( \times 10^{-8} \); nearest gene, TFCP2L1) and rs12912292 at 15q21.3 (\( P = 7.96 \times 10^{-11} \); \( P \)-meta = 1.55 \( \times 10^{-19} \); nearest gene PRTG). Case-only analyses did not reveal specific associations with TGCT histology. TFCP2L1 joins the growing list of genes located within TGCT risk loci with biologically plausible roles in developmental transcriptional regulation, further highlighting the importance of this phenomenon in TGCT oncogenesis.

INTRODUCTION

Testicular germ cell tumor (TGCT) is the most common cancer in men aged between 18 and 45, with more than 52,000 men diagnosed annually worldwide [1]. Known risk factors include a family history of the disease, a previously diagnosed germ cell tumor, subfertility, undescended testis (UDT) [2] and testicular microlithiasis [3], characterized by intratesticular calcification. Histologically, TGCT can be divided into two main...
subtypes: seminomas, which resemble undifferentiated primordial germ cells and nonseminomas, which show varying degrees of differentiation toward embryonal and extraembryonal structures. Some tumors show features of both classes (mixed histology). Both subtypes are thought to arise from progenitor germ cells through a pre-invasive phase of intratubular germ cell neoplasia (ITGCN) [4].

The cancer has a strong heritable basis, which is reflected in an observed four to eight-fold familial relative risk [5–8] and, from heritability analyses utilizing familial data, it has been estimated that genetic factors contribute to nearly half of all disease risk [9]. Despite the sizable heritable component, high penetrance TGCT susceptibility variants accounting for a sizeable proportion of genetic susceptibility have not been identified. We recently described enrichment in familial TGCT using exome sequencing of rare disruptive mutations in genes relating to ciliary and microtubule functions; however, these variants account for only a minor fraction of disease heritability [10]. In contrast, interrogation via genome-wide association studies (GWAS) for common variants of more modest effect size has proved vastly more fruitful, with a total of 50 independent risk loci proposed to date [11–22].

The two most recent, contemporaneously published TGCT GWAS reported a total of 26 novel TGCT susceptibility loci, more than doubling, from 24 to 49, the number of regions identified by preceding efforts [21, 22]. In Litchfield et al., (2017) [21], we performed a new GWAS in UK TGCT cases using the OncoArray platform (UK OncoArray study, 3,206 cases, 7,422 controls). These data were combined in meta-analysis with two previously published GWAS datasets from the UK and Scandinavia (2,313 cases, 11,633 controls), followed by replication genotyping performed for the most strongly associated loci (1,801 cases, 4,027 controls). In this analysis, in total comprising 7,319 cases and 23,082 controls, we identified 19 new loci associated with TGCT susceptibility (1p36.22, 2p13.3, 4q21.1, 4q35.2, 6q25.1, 7p14.1, 8p23.1, 11q24.2, 12p11.21, 12q15, 14q22.3, 15q22.31, 15q25.2, 16p13.11, 18p11.32, 19q11, 19q13.42, 20q13.2 and 22q11.21).

In the contemporaneously published study reported by Wang et al., (2017) [22], under the auspices of the TEsticular CAncer Consortium (TECAC), meta-analysis of data from five TGCT GWAS datasets (totaling 3,558 TGCT cases and 13,970 controls and inclusive of the UK/Scandinavian datasets used in Litchfield et al. 2017) identified associations at eight loci (2q14.2, 3q26.2, 4q35.2, 7q36.3, 10q26.13, 15q21.3, 15q22.31 and Xq28), two of which were also identified in the UK OncoArray study (4q35.2 and 15q22.31). Additional associations were also reported in the TECAC meta-analysis at previously established TGCT loci at 9p24.3 and 19p12.

In the present study, we sought independent evidence of replication at the five novel autosomal loci unique to the TECAC meta-analysis (2q14.2, 3q26.2, 7q36.3, 10q26.13, 15q21.3) using the UK OncoArray GWAS data.

RESULTS

The UK OncoArray GWAS includes data from 10,628 UK individuals, comprising 3,206 TGCT cases and 7,422 controls. The final number of SNPs passing quality control filters was 371, 504, which were used to impute genotypes at over 10 million SNPs. We looked for evidence of association for the five index SNPs reported in the TECAC meta-analysis using a frequentist approach under an additive model. We also performed meta-analysis combining data from the UK OncoArray GWAS and the TECAC meta-analysis using a fixed-effects model.

The strongest evidence for an association amongst the five loci in the UK OncoArray GWAS dataset was at 15q21.3 (Table 1; Figure 1). The reported index SNP from the TECAC meta-analysis, rs12912292, showed a highly significant association in the UK OncoArray dataset (P = 7.96 × 10−11), as did its most strongly linked directly genotyped tagging SNP (rs12899976, r2 > 1.0, D’ > 1.0, P = 2.34 × 10−15). Notably, SNPs in this region did show evidence of association in the meta-analysis undertaken in Litchfield et al., including rs12912292. However, due to poor phet and P values associated with rs12912292, an alternative SNP (rs7175728) had been chosen for replication genotyping in 1,801 cases and 4027 controls, which failed to replicate (P = 0.97, OR = 0.9986). The reported index SNP at 2q14.2, rs2713206, was not significant in the UK OncoArray dataset after correcting for multiple testing (i.e. five tests), though it was significant at a nominal threshold (P = 3.03 × 10−2; Table 1; Figure 1). Of note, a nearby directly genotyped tagging SNP in strong LD with rs2713206 (rs2713207; r2 > 0.7, D’ > 0.9) showed a stronger level of association (P = 9.44 × 10−3). For the reported index SNP at 2q14.2, the point estimate for the effect size was smaller in the UK OncoArray dataset than reported in the TECAC meta-analysis, likely a reflection of “winner’s curse”. For the reported index SNP at both loci, genome-wide significance (P < 5 × 10−8) was achieved in meta-analysis of the UK OncoArray data with the constituent TECAC datasets (Table 2).

Analysis of the UK OncoArray data did not find any evidence of association with TGCT risk for the loci at 3q26.2, 7q36.3 and 10q26.13 and did not achieve genome-wide significance when combined with the TECAC data at meta-analysis (Table 1, Table 2). Directly genotyped tagging SNPs at these three regions did not show any evidence of association with TGCT risk.

Finally, we investigated whether the two SNPs showing evidence of association in the current study (rs12912292 and rs2713206) showed differences in risk allele frequency in phenotypically-defined subgroups of TGCT cases (Table 3). Neither of the two SNPs showed
**Figure 1:** (A–B) Regional plots for loci 2q14.2 (A) and 15q21.3 (B) based on the UK OncoArray GWAS data. Triangles indicate directly genotyped SNPs while circles indicate imputed SNPs, with their position on the y axis indicating their $-\log_{10}$ association $P$ values and their position on the x axis indicating their SNP build 37 coordinates. The intensity of red shading indicates the strength of linkage disequilibrium (LD) with the index SNP (enlarged circle). Recombination rates are plotted on the z axis (light blue). The Index SNP at each locus and its mostly strongly linked tagging SNP are labelled.
Table 1: UK OncoArray GWAS data for index SNPs identified in the TECAC meta-analysis and their strongest linked tagging SNP

<table>
<thead>
<tr>
<th>Locus</th>
<th>Nearest gene</th>
<th>SNP</th>
<th>Type</th>
<th>r2</th>
<th>D’</th>
<th>Position (HG19)</th>
<th>INFO</th>
<th>Alleles</th>
<th>Case RAF</th>
<th>Control RAF</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q14.2</td>
<td>TFCP2LI</td>
<td>rs2713206</td>
<td>Imputed</td>
<td>0.97</td>
<td>C/T</td>
<td>122007941</td>
<td>0.97</td>
<td>C/T</td>
<td>0.16</td>
<td>0.15</td>
<td>1.09 (1.01–1.18)</td>
<td>3.03E-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs2713207</td>
<td>Genotyped</td>
<td>0.7</td>
<td>0.9</td>
<td>122007858</td>
<td>1.00</td>
<td>G/A</td>
<td>0.17</td>
<td>0.15</td>
<td>1.11 (1.03–1.20)</td>
<td>9.44E-03</td>
</tr>
<tr>
<td>3q26.2</td>
<td>GPR160</td>
<td>rs3755605</td>
<td>Imputed</td>
<td>0.99</td>
<td>C/T</td>
<td>169756119</td>
<td>0.99</td>
<td>C/T</td>
<td>0.40</td>
<td>0.39</td>
<td>1.05 (0.99–1.11)</td>
<td>1.32E-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs7651441</td>
<td>Genotyped</td>
<td>0.6</td>
<td>0.8</td>
<td>169738278</td>
<td>1.00</td>
<td>C/T</td>
<td>0.37</td>
<td>0.35</td>
<td>1.08 (1.01–1.15)</td>
<td>1.81E-02</td>
</tr>
<tr>
<td>7q36.3</td>
<td>NCAFPG2</td>
<td>rs11769858</td>
<td>Imputed</td>
<td>0.94</td>
<td>T/C</td>
<td>158501492</td>
<td>0.94</td>
<td>T/C</td>
<td>0.32</td>
<td>0.32</td>
<td>0.99 (0.92–1.05)</td>
<td>6.47E-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs2290393</td>
<td>Genotyped</td>
<td>0.6</td>
<td>0.9</td>
<td>158438186</td>
<td>1.00</td>
<td>G/A</td>
<td>0.39</td>
<td>0.38</td>
<td>0.97 (0.91–1.03)</td>
<td>2.81E-01</td>
</tr>
<tr>
<td>10q26.13</td>
<td>LHPP</td>
<td>rs61408740</td>
<td>Imputed</td>
<td>0.99</td>
<td>C/G</td>
<td>126274612</td>
<td>0.99</td>
<td>C/G</td>
<td>0.03</td>
<td>0.03</td>
<td>0.97 (0.81–1.16)</td>
<td>6.72E-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs1006535</td>
<td>Genotyped</td>
<td>0.2</td>
<td>0.8</td>
<td>126277624</td>
<td>1.00</td>
<td>C/T</td>
<td>0.07</td>
<td>0.07</td>
<td>0.97 (0.86–1.09)</td>
<td>5.98E-01</td>
</tr>
<tr>
<td>15q21.3</td>
<td>PRTG</td>
<td>rs12912292</td>
<td>Imputed</td>
<td>0.97</td>
<td>G/A</td>
<td>56038707</td>
<td>0.97</td>
<td>G/A</td>
<td>0.58</td>
<td>0.53</td>
<td>1.21 (1.14–1.29)</td>
<td>7.96E-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs12899976</td>
<td>Genotyped</td>
<td>1.0</td>
<td>1.0</td>
<td>55984439</td>
<td>1.00</td>
<td>A/C</td>
<td>0.57</td>
<td>0.53</td>
<td>1.22 (1.15–1.30)</td>
<td>2.34E-11</td>
</tr>
</tbody>
</table>

RAF, risk allele frequency. Risk allele in bold and underlined. INFO, information score indicating certainty of imputation (0-1). Reported P-values are from a test for trend using SNPTEST.

a significant difference in frequency between cases with seminoma (n = 1,120) compared to nonseminoma/mixed histology (n = 643), cases with testicular maldescent (n = 308) compared to those with normal descent (n = 2,837), cases with a family history of TGCT (n = 53) compared to those without (n = 3,122) or cases with unilateral (n = 3,028) compared to bilateral (n = 78) disease.

DISCUSSION

In summary, we present independent evidence supporting associations between loci at 2q14.2 and 15q21.3 and susceptibility to TGCT.

rs2713206 at 2q14.2 localizes to the intron of TFCP2LI in an LD block of ~50 kb. TFCP2LI, a member of the CP2 family of transcription factors, is a component of a complex transcriptional network involved in the establishment and maintenance of pluripotency in embryonic stem cells. TFCP2LI is highly expressed in primordial germ cells during embryogenesis [23] and is downregulated during transition of fetal gonocytes into spermatogonia [24]. TFCP2LI is not expressed in normal adult testes, though it is in intratubular germ cell neoplasia unclassified (ITGCN, formerly known as carcinoma in situ, CIS) [24], a non-invasive precursor lesion from which TGCT is widely accepted to originate. As previously reported, there are eQTL variants in LD with the index SNP. The TGCT risk allele is associated with reduced TCFP2LI expression, supporting transcriptional regulation of this gene as the functional mechanism through which the association may be mediated [22].

rs12912292 at 15q21.3 resides in a 130 kb region of LD that only contains PRTG (protogenin), which encodes an immunoglobulin superfamily transmembrane protein expressed in the developing nervous system [25]. rs12912292 displays strong eQTL effects for PRTG in muscle-skeletal (GTEx data, P = 1.9e-13) and thyroid (P = 5.1e-12) tissues; there is, however, no evidence for association of rs12912292 with expression of PRTG in either normal testes or TGCT [22].

Our data did not provide evidence supporting association with TGCT risk for SNPs at three of the loci analyzed (rs3755605 at 3q26.2, rs11769858 at 7q36.3 and rs61408740 at 10q26.13). The absence of demonstrable association for these loci in the UK OncoArray dataset could be due to power, sampling error and winner’s curse. Variable LD between genotyped marker and causal variants and/or their frequency, hidden population substructures not accounted for by principal component analysis, differences in effect modifiers and technical artefacts induced by the use of different genotyping platforms or quality-control criteria may also be contributory. More detailed and/or larger studies are required to further explore the observed differences.

We investigated whether the loci at 2q14.2 and 15q21.3 are associated with different risks in subgroups of TGCT cases characterized by specific phenotypic characteristics. Neither locus showed a significant difference in effect for the subtypes analyzed. For bilaterality, maldescent and family history, there was very limited power to detect a difference because of the small numbers examined. However, analysis of seminoma compared to nonseminoma was better powered, with > 95% power to detect effect difference of ≥ 1.5 fold; absence of difference of effects is consistent with observations for SNPs identified in earlier GWAS [11, 13] suggesting that, despite their distinct histological and biological features, these two subclasses of TGCT share a common biological pathway of oncogenesis.

TFCP2LI joins the growing list of candidate genes within TGCT risk loci linked to developmental transcriptional regulation, a key disease mechanism.
**Table 2: Meta-analysis of the UK OncoArray GWAS and the constituent TECAC datasets**

<table>
<thead>
<tr>
<th>Study</th>
<th>INFO</th>
<th>Case RAF</th>
<th>Control RAF</th>
<th>OR (95% CI)</th>
<th>(P)-value</th>
<th>(P_{\text{het}})</th>
<th>(I^2) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2713206 at 2q14.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI</td>
<td>0.92</td>
<td>0.20</td>
<td>0.15</td>
<td>1.45 (1.18–1.79)</td>
<td>4.19E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>0.98</td>
<td>0.17</td>
<td>0.15</td>
<td>1.15 (1.00–1.31)</td>
<td>4.43E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PENN</td>
<td>0.90</td>
<td>0.20</td>
<td>0.16</td>
<td>1.32 (1.06–1.64)</td>
<td>1.25E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway/Sweden</td>
<td>0.93</td>
<td>0.17</td>
<td>0.14</td>
<td>1.25 (1.09–1.44)</td>
<td>1.74E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.85</td>
<td>0.21</td>
<td>0.17</td>
<td>1.41 (0.98–2.03)</td>
<td>6.39E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK OncoArray</td>
<td>0.97</td>
<td>0.16</td>
<td>0.15</td>
<td>1.09 (1.01–1.18)</td>
<td>3.03E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
<td>1.16 (1.09–1.22)</td>
<td>3.92E-08</td>
<td>0.065</td>
<td>51.8</td>
</tr>
<tr>
<td>rs3755605 at 3q26.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI</td>
<td>0.97</td>
<td>0.44</td>
<td>0.41</td>
<td>1.14 (0.98–1.33)</td>
<td>9.31E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>0.98</td>
<td>0.43</td>
<td>0.39</td>
<td>1.21 (1.09–1.33)</td>
<td>2.25E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PENN</td>
<td>0.97</td>
<td>0.43</td>
<td>0.41</td>
<td>1.08 (0.92–1.27)</td>
<td>3.54E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway/Sweden</td>
<td>0.98</td>
<td>0.44</td>
<td>0.40</td>
<td>1.24 (1.12–1.37)</td>
<td>2.08E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.99</td>
<td>0.43</td>
<td>0.39</td>
<td>1.19 (0.91–1.55)</td>
<td>1.94E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK OncoArray</td>
<td>0.99</td>
<td>0.40</td>
<td>0.39</td>
<td>1.05 (0.99–1.11)</td>
<td>1.32E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
<td>1.11 (1.06–1.15)</td>
<td>1.10E-07</td>
<td>0.041</td>
<td>56.7</td>
</tr>
<tr>
<td>rs11769858 at 7q36.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI</td>
<td>0.93</td>
<td>0.35</td>
<td>0.33</td>
<td>1.12 (0.94–1.32)</td>
<td>1.86E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>0.95</td>
<td>0.35</td>
<td>0.31</td>
<td>1.19 (1.06–1.32)</td>
<td>1.51E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PENN</td>
<td>0.88</td>
<td>0.40</td>
<td>0.36</td>
<td>1.23 (1.04–1.47)</td>
<td>1.69E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway/Sweden</td>
<td>0.91</td>
<td>0.34</td>
<td>0.30</td>
<td>1.22 (1.10–1.37)</td>
<td>3.89E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.90</td>
<td>0.36</td>
<td>0.32</td>
<td>1.22 (0.93–1.61)</td>
<td>1.59E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK OncoArray</td>
<td>0.94</td>
<td>0.32</td>
<td>0.32</td>
<td>0.99 (0.92–1.05)</td>
<td>6.47E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
<td>1.08 (1.03–1.13)</td>
<td>1.54E-04</td>
<td>0.004</td>
<td>71.0</td>
</tr>
<tr>
<td>rs61408740 at 10q26.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI</td>
<td>0.99</td>
<td>0.04</td>
<td>0.02</td>
<td>1.68 (1.09–2.60)</td>
<td>1.89E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>0.95</td>
<td>0.04</td>
<td>0.03</td>
<td>1.64 (1.22–2.20)</td>
<td>1.05E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PENN</td>
<td>0.94</td>
<td>0.04</td>
<td>0.03</td>
<td>1.92 (1.22–3.03)</td>
<td>4.92E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway/Sweden</td>
<td>1.00</td>
<td>0.03</td>
<td>0.02</td>
<td>1.53 (1.12–2.09)</td>
<td>7.79E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.96</td>
<td>0.03</td>
<td>0.02</td>
<td>1.61 (0.69–3.76)</td>
<td>2.75E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK OncoArray</td>
<td>0.99</td>
<td>0.03</td>
<td>0.03</td>
<td>0.97 (0.81–1.16)</td>
<td>6.72E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
<td>1.15 (1.00–1.30)</td>
<td>1.27E-04</td>
<td>0.003</td>
<td>72.2</td>
</tr>
<tr>
<td>rs12912292 at 15q21.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI</td>
<td>0.95</td>
<td>0.55</td>
<td>0.51</td>
<td>1.25 (1.07–1.46)</td>
<td>4.76E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>0.99</td>
<td>0.55</td>
<td>0.53</td>
<td>1.09 (0.99–1.20)</td>
<td>9.18E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PENN</td>
<td>0.95</td>
<td>0.56</td>
<td>0.47</td>
<td>1.44 (1.23–1.70)</td>
<td>8.74E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway/Sweden</td>
<td>0.95</td>
<td>0.58</td>
<td>0.52</td>
<td>1.26 (1.14–1.39)</td>
<td>3.42E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.95</td>
<td>0.57</td>
<td>0.51</td>
<td>1.30 (1.00–1.69)</td>
<td>4.81E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK OncoArray</td>
<td>0.97</td>
<td>0.58</td>
<td>0.53</td>
<td>1.21 (1.14–1.29)</td>
<td>7.96E-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
<td></td>
<td>1.21 (1.16–1.26)</td>
<td>1.55E-19</td>
<td>0.074</td>
<td>50.2</td>
</tr>
</tbody>
</table>

RAF, risk allele frequency. Reported \(P\)-value derives from the fixed-effects inverse-variance method implemented in META.
implicated in TGCT oncogenesis [21]. Further functional evaluation is required to explore the cellular mechanisms through which the associations are mediated. The set of 50 GWAS loci identified to date are more strongly predictive of disease than the SNP sets for cancer types such as breast, colorectal and prostate cancer despite much larger GWAS in these cancer types having identified much greater numbers of SNPs: those in the highest centile for risk estimated from the TGCT SNP set have a relative risk of > 14 compared to the population risk [21, 26]. The continued success of GWAS in TGCT provides a strong rationale for continuing studies to identify additional risk loci via these methods.

**Materials and Methods**

**Sample description**

TGCT cases \( (n = 3,206) \) were ascertained via two UK studies: (1) a UK study of familial testicular cancer and (2) a systematic collection of UK TGCT cases. Case recruitment was via the UK Testicular Cancer Collaboration, a group of oncologists and surgeons treating TGCT in the UK (Supplementary Note). The studies were coordinated at the Institute of Cancer Research (ICR). Samples and information were obtained with full informed consent and Medical Research and Ethics Committee approval (MREC02/06/66 and 06/MRE06/41).

Control samples for the primary GWAS were all taken from within the UK. Specifically 2,976 cancer-free, male controls were recruited through two studies within the PRACTICAL Consortium (Supplementary Note): (1) the UK Genetic Prostate Cancer Study (UKGPCS) (age < 65), a study conducted through the Royal Marsden NHS Foundation Trust and (2) SEARCH (Study of Epidemiology & Risk Factors in Cancer), recruited via GP practices in East Anglia (2003–2009). 4,446 cancer-free female controls from across the UK were recruited via the Breast Cancer Association Consortium (BCAC).

**GWAS**

Genotyping was conducted using a custom Infinium OncoArray-500K BeadChip (OncoArray) from Illumina (Illumina, San Diego, CA, USA), comprising a 250K SNP genome-wide backbone and 250K SNP custom content selected across multiple consortia within COGS (Collaborative Oncological Gene-environment Study). OncoArray genotyping was conducted in accordance with the manufacturer’s recommendations by the Edinburgh Clinical Research Facility, Wellcome Trust CRF, Western General Hospital, Edinburgh EH4 2XU.

OncoArray data was filtered as follows: we excluded individuals with low call rate (< 95%), with abnormal autosomal heterozygosity (> 3 SD above the mean) or with > 10% non-European ancestry (based on multi-dimensional scaling); we excluded SNPs with minor allele frequency < 1%, a call rate of < 95% in cases or
controls or with a minor allele frequency of 1–5% and a call rate of < 99%; we excluded SNPs deviating from Hardy-Weinberg equilibrium ($P > 10^{-12}$ in controls and $10^{-3}$ in cases). The final number of SNPs passing quality control filters was 371,504. These data are deposited at European Genome–phenome Archive [EGA] under accession code EGAS00001001836.

Imputation

Genome-wide imputation was performed for all GWAS datasets. The 1000 genomes phase 1 data (Sept-13 release) was used as a reference panel, with haplotypes pre-phased using SHAPEIT2 [27]. Imputation was performed using IMPUTE2 software [28].

Statistical analyses

OncoArray data tests of association between imputed SNPs and TGCT was performed under an additive genetic model in SNPVTestv2.5 [29], adjusting for principal components. Inflation in the test statistics was observed at only modest levels, $\lambda_{1000} = 1.03$ [21]. The inflation factor $\lambda$ was based on the 90% least-significant SNPs [30]. The adequacy of the case-control matching and possibility of differential genotyping of cases and controls were formally evaluated using Q-Q plots of test statistics [21]. Population ancestry structure for the cohort was assessed through visualization of the first two principle components [21]; stable ancestral clustering was observed. Unadjusted case-only analyses on binary phenotypic characteristics were performed under an additive genetic model in SNPTESTv2.5, arbitrarily assigning one subdivision for a given phenotype to control status. Meta-analyses were performed using the fixed-effects inverse-variance method based on the $\beta$ estimates and standard errors from each study using META v1.7 [31]. Cochran’s Q-statistic to test for heterogeneity and the $I^2$ statistic to quantify the proportion of the total variation due to heterogeneity was calculated [32]. Regional plots were generated using visPIG software [33]. Power calculations were performed using the methods described by Skol et al. 2006 [34], implemented via the web interface at http://csg.sph.umich.edu/abecasis/cats/gas_power_calculator/index.html.

Author contributions


ACKNOWLEDGMENTS

We thank the patients with TGCT and the clinicians involved in their care for participation in this study. We thank the patients and all clinicians forming part of the UK Testicular Cancer Collaboration (UKTCC) for their participation in this study. This study would not have been possible without the contributions of M.K. Bolla (Breast Cancer Association Consortium (BCAC)), Q. Wang (BCAC), K. Michailidou (BCAC), J. Dennis (BCAC), P. Hall (Collaborative Oncological Gene-environment Study (COGS)); D.F. Easton (BCAC), A. Berchuck (Ovarian Cancer Association Consortium), R. Eeles (PRACTICAL), G. Chenevix-Trench (The Consortium of Investigators of Modifiers of BRCA1/2), J. Dennis, P. Pharoah, A. Dunning, K. Muir, J. Peto, A. Lee and E. Dicks. We also thank the following for their contributions to this project: J. Simard, P. Kraft, C. Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory; and K.F. Doheny and the staff of the Center for Inherited Disease Research (CIDR) genotyping facility. The results published here are based in part on data generated by the TCGA Research Network. This study makes use of data generated by the Wellcome Trust Case Control Consortium 2 (WTCCC2). We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

GRANT SUPPORT

We acknowledge National Health Service funding to the National Institute for Health Research Biomedical Research Centre. We thank the UK Genetics of Prostate Cancer Study (UKGPCS) study teams for the recruitment of the UKGPCS controls. Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) (U19 CA 148537 for Elucidating Loci Involved in Prostate cancer Susceptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN2682012000081). Additional analytical support was provided by NIH NCI U01 CA188392. The PRACTICAL consortium was supported by Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A10118.
REFERENCES


