Susceptibility to testicular germ cell tumors (TGCT) has a significant heritable component, and genome-wide association studies (GWASs) have identified association with variants in several genes, including KITLG, SPRY4, BAK1, TERT, DMRT1 and ATF7IP. In our GWAS, we genotyped 349 TGCT cases and 919 controls and replicated top hits in an independent set of 439 cases and 960 controls in an attempt to find novel TGCT susceptibility loci. We identified a second marker (rs7040024) in the doublesex and mab-3-related transcription factor 1 (DMRT1) gene that is independent of the previously described risk allele (rs755383) at this locus. In combined analysis that mutually conditions on both DMRT1 single nucleotide polymorphism markers, TGCT cases had elevated odds of carriage of the rs7040024 major A allele [per-allele odds ratio (OR) = 1.48, 95% confidence interval (CI) 1.23, 1.78; P = 2.52 x 10^-5] compared with controls, while the association with rs755383 persisted (per allele OR = 1.26, 95% CI 1.08, 1.47, P = 0.0036). In similar analyses, the association of rs7040024 among men with seminomatous tumors did not differ from that among men with non-seminomatous tumors. In combination with KITLG, the strongest TGCT susceptibility locus found to date, men with TGCT had greatly elevated odds (OR = 14.1, 95% CI 5.12, 38.6; P = 2.98 x 10^-7) of being double homozygotes for the risk (major) alleles at DMRT (rs7040024) and KITLG (rs4474514) when compared with men without TGCT. Our findings continue to corroborate that genes influencing male germ cell development and differentiation have emerged as the major players in inherited TGCT susceptibility.

*To whom correspondence should be addressed at: University of Pennsylvania, 219 Blockley Hall, Philadelphia, PA 19104. Tel: +1 2155733282; Fax: +1 2155731050; Email: pkanetsk@mail.med.upenn.edu

© The Author 2011. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
INTRODUCTION

In the USA, testicular germ cell tumors (TGCT [MIM 273300]) are the most common cancers in young men, with a peak incidence among those aged 25–34 years. The incidence of TGCT among white men in the USA has increased substantially over the past 30 years from 4.1 per 100 000 in 1975 to 7.0 per 100 000 in 2007 (1), and similar increases are seen worldwide (2). The disease incidence varies widely across racial groups, with a 5-fold lower rate among black men in the USA (1). Differences in incidence of TGCT also exist across countries and continents, ranging from Denmark (9.2 per 100 000) to Algeria (0.2 per 100 000), which is consistent with racial differences and lower rates in non-white groups (3).

Familial aggregation of TGCT has been documented since the 1930s, and family history is the strongest known risk factor for these malignancies (4,5). Risk of TGCT repeatedly has been shown to be increased among first-degree relatives of affected men, with risk to brothers (estimates range from 5- to 19-fold) being stronger than that to fathers (estimates range from 2- to 4-fold) (6–14). Both mono- and dizygotic twins of affected men have increased risk of TGCT (15,16). Genetic effects have been estimated to account for 25% of the familial aggregation along with the top observed marker (rs755383), which was previously showed to be associated with TGCT by Turnbull et al. (20), was taken into replication along with the top observed marker (rs7040024), both of which exceeded the $P < 5 \times 10^{-6}$ threshold (Supplementary Material, Fig. S2).

We did not replicate associations with rs2279070 ($P_{\text{trend}} = 0.62$) and rs4834214 ($P_{\text{trend}} = 0.54$) at 4q28.2 or rs6951213 ($P_{\text{trend}} = 0.48$) and rs10281060 ($P_{\text{trend}} = 0.68$) at PFTK1. However, we observed statistically significant associations with rs7040024 ($P_{\text{trend}} = 2.09 \times 10^{-6}$) and rs755383 ($P_{\text{trend}} = 2.08 \times 10^{-6}$) at DMRT1. Using a combined set of discovery and replication samples, TGCT cases had greater odds of carriage of the major T allele in rs7040024 than controls ($P = 1.41 \times 10^{-11}$; odds ratio (OR) = 1.70, 95% confidence interval (CI) 1.46, 1.99) and greater odds of carriage of the major T allele in rs755383 ($P = 8.61 \times 10^{-10}$; OR = 1.50, 95% CI 1.32, 1.7) (Table 2). In addition to the case–control analysis, we performed a case–parent analysis in 179 triads and 135 dyads that showed homozygous carriage of the A allele in rs7040024 than controls ($P = 0.0033$; relative risk = 3.42, 95% CI 1.50, 7.76) as was homozygous carriage of the T allele in rs755383 ($P = 1.20 \times 10^{-10}$; relative risk = 4.67, 95% CI 2.92, 7.46). These analyses support the finding of an association between the variation at the DMRT1 locus and TGCT susceptibility.

RESULTS

Information on age, risk factors and tumor characteristics for the discovery and replication sample sets are given in Table 1. The calculated genomic control inflation ($\lambda$) factor in the discovery set was 0.942, and hence we report unadjusted test statistics (21,22). We again noted our previously reported statistically significant associations with markers at 12q22 in KITLG ($P < 5.0 \times 10^{-8}$), 2p14 and 5q13.3 near SPRY4 ($P < 5.0 \times 10^{-6}$; Fig. 1 and Supplementary Material, Table S1) (18). Six additional markers at five additional autosomal loci also were associated with TGCT ($P < 5.0 \times 10^{-8}$). To screen out hits representing likely false positive associations, we imputed genotypes in the surrounding regions of these six markers using data from 1000 genomes (March 2010 release) (Fig. 2A and Supplementary Material, Fig. S1). After imputation, neither observed nor imputed markers at 8q21.3 and 13q12.3 maintained significance at $P < 5 \times 10^{-6}$ and were excluded from further study, whereas markers at 4q28.2, 7q21.13 and 9p24.3 continued to exceed the $P < 5 \times 10^{-6}$ threshold. The selected markers on chromosomes 4 and 7 that were brought forward into replication included two imputed (rs2279070, rs4834214) that mapped to 4q28.2, as the observed single nucleotide polymorphism (SNP) could not be designed for replication genotyping, and two observed (rs6951213, rs10281060) that mapped within introns 1 and 2 of PFTK1 on 7q21.13. For chromosome 9 markers, the top three imputed markers using 1000 genomes were incompatible with our replication genotyping platform, so we re-imputed markers in this region using data from HapMap (release 22). The top imputed marker (rs755383), which was previously showed to be associated with TGCT by Turnbull et al. (20), was taken into replication along with the top observed marker (rs7040024), both of which exceeded the $P < 5 \times 10^{-6}$ threshold (Supplementary Material, Fig. S2).

Analyses that additionally adjusted for cryptorchidism or family history of TGCT were limited to the replication sample because information on risk factors was not collected...
for discovery phase controls. Here, associations with both DMRT1 markers (simultaneously considered) and TGCT were unchanged following further adjustment for either cryptorchidism or family history. Among men without a family history of TGCT or men without cryptorchidism, the ORs were comparable with overall results. This finding suggests that the effects of DMRT1 variants are not based on the pathologic mechanisms leading to these known strong TGCT risk factors.

In the combined set, we explored whether the effects of DMRT1 rs7040024 and rs755383 were modified by genotypes at KITLG rs4474514 and SPRY4 rs6897876, loci we previously reported as independently associated with TGCT status (18). In models that assessed the joint effects of DMRT1 genotypes with KITLG and SPRY4 genotypes, we failed to detect departures from multiplicativity. However, we did observe statistical evidence supporting departure from additivity (i.e. additive synergy) for DMRT1 rs7040024 and KITLG rs4474514 (synergy index, S = 2.2, 95% CI 1.1, 4.2; P = 0.02). That is to say, the OR associated with homozygous carriage of the risk allele at both loci (OR = 14.1, 95% CI 5.12, 38.6; P = 2.98 × 10^{-2}) was greater than the additive effect of the OR for homozygous carriage of the risk allele at either locus alone (OR = 2.44, 95% CI 0.62, 9.56; P = 0.20 for DMRT1 rs7040024; OR = 5.59, 95% CI 1.84, 7.0, P = 2.45 × 10^{-3} for KITLG rs4474514; Fig. 3). Here, the comparison group was individuals who carried a total of zero or one risk allele at both loci (n = 86), which was necessary because only five controls and no TGCT cases were doubly homozygous for the non-risk alleles. We did not find evidence of departure from additivity for DMRT1 rs7040024 and SPRY4 rs6897876 (synergy index, S = 2.2, 95% CI 0.7, 7.2, P = 0.20).

Evidence of association with markers at ATF7IP, BAK1 and TERT was inconsistent across our sample sets. Although TGCT status was associated with rs11055991 (ATF7IP), rs210138 (BAK1), rs444697 (BAK1) and rs2736100 (TERT) in our discovery set, no associations were noted in our replication set (Supplementary Material, Table S2). In contrast to reported findings by Turnbull et al. (20), we did not find associations at TERT to be stronger in tumors of the seminoma subtype.

**DISCUSSION**

We verified the finding that variation at 9p24.3 within DMRT1 is associated with TGCT susceptibility (20) and report the
identification of a second risk allele (rs7040024) at this locus that is independent of the one previously reported (rs755383). In combined analysis and after mutual adjustment, TGCT cases had 50% greater odds of carriage per major A allele at rs7040024, while simultaneously having nearly 25% greater odds of carriage per major T allele at rs755383. Of note, rs755383 was the most statistically significant (imputed) marker at this locus and the marker genotyped directly by Turnbull et al. We anticipated that these markers might be independently associated with TGCT because of their low-to-moderate linkage disequilibrium ($r^2 = 0.29$) with each other, and the fact that these two markers do not reside in the same haplotype block. Thus, we have identified a second SNP within DMRT1 that is associated more strongly

Table 2. Associations of TGCT with replicated DMRT1 SNP markers

<table>
<thead>
<tr>
<th>Analysis group</th>
<th>Gene marker&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Genotype count&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Phase</th>
<th>OR (95% CI)</th>
<th>Heterozygote&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Homozygote&lt;sup&gt;d&lt;/sup&gt;</th>
<th>P-value&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>rs7040024 A/C</td>
<td>493/356/357</td>
<td>Cases</td>
<td>Discovery</td>
<td>1.71 (1.37, 2.14)</td>
<td>1.74 (0.89, 3.41)</td>
<td>2.96 (1.54, 5.70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>276/110/11</td>
<td></td>
<td>Replication</td>
<td>1.69 (1.36, 2.10)</td>
<td>1.77 (0.89, 3.49)</td>
<td>2.96 (1.52, 5.73)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>329/389/144</td>
<td></td>
<td>Combined</td>
<td>1.70 (1.46, 1.99)</td>
<td>1.75 (1.08, 2.83)</td>
<td>2.96 (1.86, 4.71)</td>
</tr>
<tr>
<td></td>
<td>rs755383&lt;sup&gt;f&lt;/sup&gt; T/C</td>
<td>371/415/133</td>
<td></td>
<td>Discovery</td>
<td>1.48 (1.23, 1.78)</td>
<td>1.39 (0.83, 2.31)</td>
<td>2.00 (1.25, 3.52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>185/172/38</td>
<td></td>
<td>Replication</td>
<td>1.64 (1.35, 1.98)</td>
<td>1.71 (1.07, 2.73)</td>
<td>2.74 (1.73, 4.35)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>329/389/144</td>
<td></td>
<td>Combined</td>
<td>1.50 (1.32, 1.71)</td>
<td>1.68 (1.24, 2.28)</td>
<td>2.13 (1.43, 3.18)</td>
</tr>
<tr>
<td></td>
<td>rs755383&lt;sup&gt;f&lt;/sup&gt; T/C</td>
<td>120/36/3</td>
<td></td>
<td>Discovery</td>
<td>1.39 (1.06, 1.73)</td>
<td>1.59 (0.71, 2.72)</td>
<td>1.87 (0.95, 3.69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>108/104/26</td>
<td></td>
<td>Replication</td>
<td>1.54 (1.14, 2.09)</td>
<td>1.17 (0.58, 2.34)</td>
<td>2.02 (1.03, 3.95)</td>
</tr>
<tr>
<td>Seminomatous</td>
<td>rs7040024 A/C</td>
<td>78/31/4</td>
<td>Cases</td>
<td>Discovery</td>
<td>1.46 (1.13, 1.88)</td>
<td>1.66 (0.76, 3.53)</td>
<td>2.35 (1.10, 5.04)</td>
</tr>
<tr>
<td>tumors</td>
<td></td>
<td>156/74/8</td>
<td></td>
<td>Replication</td>
<td>1.54 (1.25, 1.89)</td>
<td>1.59 (0.84, 2.98)</td>
<td>2.42 (1.31, 4.46)</td>
</tr>
<tr>
<td>Non-seminomatous</td>
<td>rs7040024 A/C</td>
<td>154/67/6</td>
<td>Cases</td>
<td>Discovery</td>
<td>1.35 (1.06, 1.67)</td>
<td>1.39 (0.71, 2.32)</td>
<td>1.87 (0.95, 3.69)</td>
</tr>
<tr>
<td>tumors</td>
<td></td>
<td>120/36/3</td>
<td></td>
<td>Replication</td>
<td>1.39 (1.17, 1.65)</td>
<td>1.37 (0.93, 2.03)</td>
<td>1.91 (1.30, 2.80)</td>
</tr>
<tr>
<td></td>
<td>rs755383&lt;sup&gt;f&lt;/sup&gt; T/C</td>
<td>126/90/13</td>
<td>Cases</td>
<td>Discovery</td>
<td>1.54 (1.14, 2.09)</td>
<td>1.17 (0.58, 2.34)</td>
<td>2.02 (1.03, 3.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>77/68/12</td>
<td></td>
<td>Replication</td>
<td>1.39 (1.17, 1.65)</td>
<td>1.37 (0.93, 2.03)</td>
<td>1.91 (1.30, 2.80)</td>
</tr>
</tbody>
</table>

<sup>a</sup>dbSNP rs number and risk/non-risk alleles.

<sup>b</sup>Number of individuals genotyped as homozygous for the risk allele/heterozygous for the risk allele/homozygous for the non-risk allele. MAF for discovery phase markers given in Supplementary Material, Table S1.

<sup>c</sup>OR for heterozygous carriage of risk allele compared with homozygous carriage of non-risk allele.

<sup>d</sup>OR for homozygous carriage of risk allele compared with homozygous carriage of non-risk allele.

<sup>e</sup>Test for trend.

<sup>f</sup>OR, 95% CI and P-value determined from a logistic regression model of combined data containing main effects of both DMRT1 markers.

<sup>g</sup>Genotype counts, OR and 95% CI for discovery phase rs755383 estimated from data imputed using HapMap (release 22).
with TGCT susceptibility than that previously reported. Our finding, together with others recently identified for TGCT susceptibility loci, begins to unravel the complex architecture of inherited genetics of TGCT.

We acknowledge that the use of a comparison group consisting of men who were undergoing routine cardiac catheterization, of whom 76% \( (n = 700) \) had angiographically confirmed coronary artery disease (CAD), may have impacted our findings to the degree that there are common genetic factors linking CAD and TGCT. To explore this possibility, we compared genotype frequencies of markers at \( BAK1 \), \( DMRT1 \) and \( TERT \) (Table 2 and Supplementary Material, Table S2) among controls with and without frank CAD and noted no statistically significant differences \( (0.09 \leq P_{\text{corr}} \leq 0.64 \text{ for all comparisons}) \). Further, we compared the observed minor allele frequencies (MAFs) at these markers among CAD controls to those observed in our replication-phase controls, HapMap CEU samples and Coriell CEPH samples. We observed slight fluctuations in allele frequencies among these four groups ranging from a 3 to 8% difference; however, for each marker, the allele frequency among the CAD controls was consistent with that in at least one of the three population-based sample sets. These comparisons lead us to believe that little bias was introduced by using CAD controls. As well, we recognize that 90% of CAD controls were 46 years or older and had already passed the peak age of TGCT development. Based on available age-specific TGCT rates, we estimated that only four TGCT cases would be expected to have arisen in this control group \( (1) \). It is unlikely that this potential small misclassification of phenotype would have biased results appreciably.

\( DMRT1 \) is a member of a zinc finger-like DNA-binding motif (DM domain) gene family. Genes with the DM domain are highly conserved and play crucial roles in male development and sex determination across the phylogenetic spectrum from flies and nematodes to birds and vertebrates \( (23) \). \( Dmrt1 \) in mice is expressed only in the gonad and is essential for postnatal germ line maintenance and differentiation of germ cells, specifically radial migration, mitotic reactivation and survival \( (24,25) \). \( Dmrt1 \) is required for normal maturation of Sertoli cells; the cells fail to polarize and stop proliferating when this gene is knocked out \( (24) \). Increased dosage of \( DMRT1 \) facilitates male development, and decreasing gene dosage leads to feminization of the gonads. In humans, deletion of the region on 9p-containing \( DMRT1 \) leads to male-to-female sex reversal and is associated with the development of gonadoblastoma \( (26,27) \). \( Dmrt1 \) \((\text{null})\) mice also have severely dysgenetic testes, resembling those seen in XY individuals with the loss of 9p, and fail to undergo normal germ cell development, dying by P14 \( (24,28) \). Moreover, a loss of \( Dmrt1 \) also is associated with a high rate of teratomas in 129Sv mice with 90% of double knockout mice developing tumors \( (29) \). In our data, variants at \( DMRT1 \) may be slightly more strongly associated with non-seminomatous germ cell tumors, of which teratoma is one type, than with seminomas; but the difference between tumor types was not statistically significant. We do not believe our inability to detect a statistical difference in the association between common \( DMRT1 \) variants by histological tumor type should be interpreted as contradicting the evidence from model systems for a role of this gene in human teratoma susceptibility. Rather, the sample size was modest, and consequently there was limited power to detect differences between tumor types.

TGCT is believed to arise from undifferentiated primordial germ cells, which progress to the precursor lesion of TGCT-termed intratubular germ cell neoplasia undifferentiated \( (\text{ITGCNU}) \) \( (30–33) \). All ITGCNU is thought to advance to frank TGCT, i.e. there is no spontaneous regression. Consistent with this hypothesis, predisposition to TGCT is increased among patients with different types of disorders of sexual development and delayed differentiation of primordial germ cells \( (34) \). The data from the TGCT GWA studies are strongly supportive of this model of disease development, as male germ cell differentiation is the biological pathway linking \( DMRT1 \) to the TGCT susceptibility loci \( KITLG \), \( SPRY4 \) and \( BAK1 \).

TGCT susceptibility has been linked to infertility or subfertility in several studies, and the majority of evidence points to common etiologic risk factors for these conditions. The most convincing studies have followed men with known semen quality for outcomes including TGCT and have examined fertility in brothers of men with TGCT \( (35–39) \). Interestingly, \( Dmrt1 \) is essential for fertility in mice \( (24) \). Thus, the association between common variation at \( DMRT1 \) and

---

**Figure 3.** Joint effects of \( DMRT1 \) genotypes with \( KITLG \) and \( SPRY4 \) genotypes. OR and 95% CI for main effects adjusted for the second SNP marker are given in the row and column headers. Within cells, frequency count (control/case) and OR (95% CI) for joint genotypes are given with referent genotype(s) represented as shaded cells.
TGCT provides further support to the notion of shared risk factors for infertility and these malignancies. Rapley et al. (19) and Turnbull et al. (20) also found evidence in their discovery and replication sets of association of TGCT with rs2900333 and rs4346018 at the ATF7IP locus, with rs210138 in the BAK1 locus and with rs2736100 and rs4635969 in the TERT locus. We genotyped TERT rs31489 (which reached genome-wide significance in the UK study, but was not brought forth into replication) in our replication sample set and looked up results in our discovery set for three other TERT markers in common with the UK study: rs2736100, rs4016810 (which is in linkage disequilibrium with rs4635969, \( r^2 = 0.67 \)) and rs4975616. We genotyped BAK1 rs210138 in our replication sample and looked up results for this and two other BAK1 markers reaching genome-wide significance in the UK study: rs210139 and rs444697 (which is in linkage disequilibrium with rs375555, \( r^2 = 0.72 \) and rs4713646, \( r^2 = 0.97 \)). At ATF7IP, we genotyped rs2900333 and rs4346018 in our replication sample and looked up results in our discovery set for rs11055991, which was in linkage disequilibrium with both rs2900333 (\( r^2 = 0.72 \)) and rs4346018 (\( r^2 = 0.95 \)). While we were unable to verify previous finding in our replication sample set, in our discovery set we found some evidence of association with ATF7IP rs11055991 (\( P = 0.052 \)), BAK1 rs210138 (\( P = 0.0074 \)), BAK1 rs444697 (\( P = 0.028 \)) and TERT rs2736100 (\( P = 0.020 \)) and after adjustment for the false discovery rate (40) observed the following \( P \)-values: 0.041, 0.041, 0.077 and 0.073, respectively.

Our inability to verify the associations at TERT and ATF7IP may be due to a difference in power of the UK study and the current study. Given the size of our discovery or replication data sets, and assuming \( \alpha = 0.05 \) (uncorrected for multiple comparisons), \( 1 - \beta = 0.80 \), an additive genetic model, and a range of MAF from 40 to 50% that correspond to observed MAF for markers in TERT and ATF7IP, the minimal detectable per allele OR ranged from 1.29 to 1.26 (41), which is greater than those observed at these loci in our study. As the white non-Hispanic population in the USA is more genetically heterogeneous than the white UK population, the associations with specific alleles in our study may be attenuated, exacerbating decreased power.

Only 3% of the TGCT cases did not carry a risk (major) allele at DMRT1 rs7040024; this proportion decreased to 1% for KITLG. Considering these loci jointly, we observed no case who was a double homozygote for the non-risk (minor) alleles at these loci. This finding suggests that carriage of at least one and potentially both major alleles may be necessary for TGCT development. Men with TGCT are 14 times more likely to be doubly homozygous for the risk alleles of KITLG and DMRT1 than those who carry a total of one or zero risk alleles. If associations with genotypes at these loci are further validated, they may identify a group of men at increased risk of developing TGCT; thus in the future, these findings may have implications for screening. At present, screening for TGCT in the general population or in men at increased risk for TGCT such as those with cryptorchidism or a family history of TGCT is not recommended by the US Preventative Services Health Task Force at present (42,43).

The relatively small number of individuals in our replication set with these conditions prevented us from assessing the extent to which the risk alleles influenced TGCT development among this subset. Future pooled analyses will increase power for such important analyses.

While in large part, the per-allele ORs identified for TGCT susceptibility alleles are stronger than those identified for other cancers; these tumors are rare and the studies are considerably smaller than GWA studies of many other cancers. Thus, it is likely that further genes remain to be identified, potentially through pooled or meta-analysis (44). Based on these and others’ similar findings, variation in pathways that influence male germ cell development and differentiation, including KITLG, SPRY4, BAK1 and DMRT1, is emerging as the major player in inherited TGCT susceptibility. Still, the biological mechanism and functional impact of the true risk alleles to which our observed risk alleles may be linked will need to be determined in order to better comprehend the genetic basis of TGCT. Other genes that play crucial roles in differentiation of the male gonad or germ cell development, such as SOX9 and STRA8, should be considered potential candidate genes. The identification of this pathway as important in TGCT susceptibility supports the hypothesis that TGCT develops from undifferentiated primordial germ cells and gives further biological basis for the known association between infertility and TGCT.

**MATERIALS AND METHODS**

**Ethics statement**

For both the GWA and replication studies, each participant provided written informed consent approved by their local Institutional Review Boards.

**Genome-wide association study**

In the discovery phase, we included 455 TGCT patients, most of whom were approached while seeking care at the University of Pennsylvania Health System or Fox Chase Cancer Center. TGCT cases not identified through this hospital-based mechanism were identified through the Pennsylvania State Cancer Registry and contacted by mail to solicit for study participation. All TGCT patients were asked to complete a self-administered questionnaire that elicits information on known and presumptive risk factors for TGCT and to provide a biospecimen. Each patient is classified as having seminomatous or non-seminomatous (including yolk sac, choriocarcinoma, embryonal, teratoma and mixed cell, i.e. having both non-seminomatous and seminomatous aspects) TGCT based on histological diagnosis. We obtain this information through medical record review for those participants recruited in person and directly from the Cancer Registry for participants recruited via mail. Only participants with primary disease in the testis are included.

Controls were selected from PennCATH, a multi-institutional hospital-based study of angiographic CAD in almost 4000 subjects undergoing cardiac catheterization. This study investigates the association of biochemical and genetic factors for CAD and its risk factors (45); information on personal history of cancer was not collected. For our
comparison group, we selected only males enrolled in PennCATH from the Philadelphia region (n = 932) independent of their disease status.

We used the Affymetrix® Genome-Wide Human SNP Array 6.0 to obtain genotypes for TGCT cases. We used the Birdseed algorithm to determine genotypes for the combined TGCT case and CAD control sample set (46). Among the 455 case samples, 19 subsequently were excluded for not meeting case eligibility (2 Leydig cell tumors, 1 female germ cell tumor erroneously coded as TGCT, 16 non-TGCT samples) and 11 replicate samples with lower genotyping call rates were excluded. Of the 425 unique samples from TGCT cases, 48 (11%) were excluded because of a low (<95%) genotyping call rate. Seven (1.5%) men were excluded because of lower than expected genotypic heterozygosity across called markers (FST ≥ 0.06), and 21 (6.2%) because of non-European ancestry as determined by multidimensional scaling (47); no cases were excluded for cryptic relatedness (proportion of genotypes IBD for all cases was <0.20). The resulting 349 samples represent a 26% increase in the number of TGCT cases analyzed in this report compared with that (n = 277) used in the original analysis. Controls had been genotyped previously using the Affymetrix® Genome-Wide Human SNP Array 6.0 platform and had passed the same genotyping quality measures used for TGCT cases. Among the 932 CAD controls, 13 were excluded because of female or ambiguous sex on genotyping.

After excluding 82 981 (11.3%) markers with a MAF (in the total sample) < 0.05; 1872 (0.3%) that deviated from Hardy–Weinberg equilibrium (P < 1 × 10−7); 38 060 (5.2%) with an individual genotype call rate < 0.95 and 233 (0.03%) invalid markers, 609 482 markers remained in the discovery phase.

Replication study
To independently replicate findings of the discovery phase, we used 439 cases and 960 controls from a population-based case–control study of TGCT in western Washington State. These numbers represent an 18% increase in cases and 12% increase in controls compared with those (n = 371 and n = 860, respectively) used in the original analysis. Briefly, all cases had first primary TGCT diagnosed between 1999 and 2008 and were residents of three urban counties of western Washington aged 18–44 years at diagnosis. Control subjects did not have a personal history of TGCT and were frequency matched on age and ascertainment from the general population of the three counties using random digit telephone dialing. Family history of TGCT among first-degree relatives and personal history of cryptorchidism were ascertained by questionnaire. Only cases and controls who self-identified as white, non-Hispanic were included in the replication study. In addition, 314 sets of mothers and fathers of cases (179 case-triads and 135 case-dyads) were used for the case–parent analysis. Methods for recruitment of TGCT cases and parents in this study previously have been published (48).

Genotyping
Replication genotyping was accomplished using the iPLEX Mass Array platform (Sequenom, Inc.) according to the manufacturer’s specifications. Genotyping was run in duplicate for 1248 marker pairs (208 sample pairs per each of the six markers in replication). Genotyping calls were made without knowledge of case or duplicate status. In total, 26 (2.1%) calls were discordant, ranging from 0.5 (rs7040024, rs6951213) to 5.3% (rs2279070); the Spearman correlation coefficient was 0.98 (ranging from 0.90 for rs2279070 to >0.99 for rs7040024). We also re-genotyped ~80% of TGCT cases from the discovery phase for the three observed markers taken into replication. The Spearman correlation coefficient between genotype calls obtained from the Affymetrix® chip and iPLEX platform for these three markers ranged from 0.96 to 0.97.

We also genotyped in our replication sample at least one top hit in BAK1, TERT and ATF7IP, which were previously observed to be associated with TGCT susceptibility (19,20). Where possible, we also looked up associations in our discovery sample for common markers at these loci.

Statistical analysis
For the discovery phase, we used PLINK software to calculate rates of heterozygosity, assess population stratification using multi-dimensional scaling methods and test markers for Hardy–Weinberg equilibrium (49). PLINK also was used to determine allelic associations among the 349 TGCT cases and 919 CAD controls and to assess their statistical significance by Fisher’s exact test. For top hits, we determined ORs and 95% CIs for the per allele, heterozygous and homozygous carriage of risk alleles with case status (overall and among specified subgroups). Trend across genotype categories was assessed by the Cochran–Armitage test for trend. We used multinomial logistic models to obtain simultaneously the OR and 95% CI for the association between TGCT subtypes defined on the basis of histology. To determine whether the observed association with a given SNP marker was independent of the effects of other SNP markers, we used logistic regression models that included the main effects of each SNP marker. To investigate the
joint effect of two SNP markers on TGCT, an indicator variable was coded to represent the various combinations of genotypes at both markers, and this indicator variable was then entered into a logistic regression model as a class variable. The synergy index and its 95% CI were used to assess departure from additivity of genetic effects (53). All analyses using the combined data set incorporated an indicator variable for study phase.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We would like to thank both the men with and without TGCT and parents of TGCT patients, who contributed to the study. We thank D. Pucci, K. Robertson, E. Ellis Ohr, A. Mackley, A. Tran, M. Shellenger, C. Panks, K. Leach and D. Jacke for patient recruitment and database maintenance. We thank S. Fish for her expert assistance with DNA extraction and genotyping. We would like to thank Drs S. Lesko and G. Bunin, consultants for the ongoing TGCT project at Penn. We acknowledge Drs C. Turnbull and N. Rahman for providing data from their GWA study of TGCT pre-publication. The Pennsylvania Cancer Registry data were supplied by the Bureau of Health Statistics & Research, Pennsylvania Department of Health, Harrisburg, Pennsylvania. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations or conclusions.

Conflict of Interest statement. None declared.

FUNDING

The work was supported by the Abramson Cancer Center at the University of Pennsylvania and National Institute of Health grants (R01CA114478 to P.A.K., K.L.N.; R01CA085914 to C.C., J.R.S., S.M.S.; Fox Chase Cancer Center Support Grant P30 CA06927; and Cancer Biostatistics Training Grant CA093283 to S.V.). Recruitment of controls from the University of Pennsylvania Health System was supported by the Cardiovascular Institute of the University of Pennsylvania, and genotyping of controls was supported by GlaxoSmithKline through an Alternate Drug Discovery Initiative research alliance award (to M.P.R. and D.J.R.) with the University of Pennsylvania School of Medicine. Services provided in the Penn Molecular Diagnosis and Genotyping Facility are supported by the Abramson Cancer Center core grant 5P30CA016520.

REFERENCES


