ORIGINAL ARTICLE

Thyroid cancer susceptibility polymorphisms: confirmation of loci on chromosomes 9q22 and 14q13, validation of a recessive 8q24 locus and failure to replicate a locus on 5q24

Angela M Jones, Kimberley M Howarth, Lynn Martin, Maggie Gorman, Radu Mihai, Laura Moss, Adam Auton, Catherine Lemon, Hisham Mehanna, Hosahalli Mohan, Susan E M Clarke, Jonathan Wadsley, Elena Macias, Andrew Coatesworth, Matthew Beasley, Tom Roques, Craig Martin, Paul Ryan, Georgina Gerrard, Danielle Power, Caroline Bremmer, Luis G Carvajal-Carmona

The TCUKIN Consortium,* Ian Tomlinson,1,16 Luis G Carvajal-Carmona1

ABSTRACT

Five single nucleotide polymorphisms (SNPs) associated with thyroid cancer (TC) risk have been reported: rs2910164 (5q24); rs9683267 (8q24); rs965513 and rs1867277 (9q22); and rs944289 (14q13). Most of these associations have not been replicated in independent populations and the combined effects of the SNPs on risk have not been examined. This study genotyped the five TC SNPs in 781 patients recruited through the TCUKIN study. Genotype data from 6122 controls were obtained from the CORGI and Wellcome Trust Case-Control Consortium studies. Significant associations were detected between TC and rs965513A (p = 6.35 × 10−34), rs1867277A (p = 5.90 × 10−24), rs944289T (p = 6.95 × 10−7), and rs9683267G (p = 0.016). rs9683267 was most strongly associated under a recessive model (pG vs GT + TT = 0.004), in contrast to the association of this SNP with other cancer types. However, no evidence was found of an association between rs2910164 and disease under any risk model (p > 0.7). The rs1867277A association remained significant (p = 0.008) after accounting for genotypes at the nearby rs965513 (p = 2.3 × 10−13) and these SNPs did not tag a single high risk haplotype. The four validated TC SNPs accounted for a relatively large proportion (~11%) of the sibling relative risk of TC, principally owing to the large effect size of rs965513 (OR 1.74).

INTRODUCTION

Thyroid cancer (TC) is the most common endocrine malignancy and a complex disease with a largely unknown aetiology.1 TC is characterised by one of the strongest familial relative risks in cancer. First degree relatives of TC patients are up to 8.6 times more likely to develop TC than the general population.2 Most of the genetic variation associated with TC remains uncharacterised, and it is likely to be explained by variants of moderate or low penetrance.

A number of recent studies have identified single nucleotide polymorphisms (SNPs) associated with TC risk on chromosomes 5q24, 8q24, 9q22, and 14q13.3–6 Two of these SNPs, rs965513 (9q22) and rs944289 (14q13), were found through a multi-stage, genome-wide association (GWA) study in the Icelandic population.7 Subsequent replication of association was found in smaller sample sets from Ohio, USA and Spain. The other three TC SNPs were discovered through candidate gene or SNP approaches.4–6 rs2910164 (5q24) was chosen because it lay within pre-miR-146a, a microRNA upregulated in TC. An association with TC was found in samples from Poland, Finland and Ohio, and the C allele at rs2910164 was found to decrease levels of pre- and mature miR-146A.4 rs9683267 (8q24) is associated with the risk of several cancers, including those of the prostate, colon and ovary, and was assessed as a TC SNP for this reason. It showed a borderline significant association with TC in the Polish population.6 rs1867277 was studied because it lies in the 5′ UTR of FOXE1 (or Thyroid Transcription Factor 2), a key gene involved in thyroid organogenesis.8 rs1867277 and rs965513 are in moderate pairwise linkage disequilibrium (LD) in Europeans (r2 = 0.39, D′ = 0.73, http://www.1000genomes.org/); rs1867277 was strongly associated with TC risk in Spanish and Italian cohorts.5 None of the three candidate SNP associations has been replicated in independent studies.

The aim of this study was to examine these five TC SNPs in a relatively large UK case–control sample set, to validate or refute their associations with TC in this population, and to estimate the proportion of the familial risk of TC for which they account.

PATIENTS AND METHODS

Study samples

We recruited 781 white UK patients of northern European origin with histologically confirmed non-medullary TC through the Thyroid Cancer Genetic investigation in the UK (TCUKIN) study. In addition to obtaining standard clinicopathological information from medical records and a questionnaire completed by each patient, the
participants donated a blood sample which was used to isolate genomic DNA. The Southampton and South West Hampshire Research Ethics Committee (A) approved the TCUKIN protocol.

**SNP genotyping and control genotype data**

We genotyped the TCUKIN samples at the five SNPs (rs2910164, rs6985267, rs695513, rs1867277, rs944289) using the KAspar system. Probes used to genotype these polymorphisms are shown in supplementary table 1. For comparison, we used available genotype data from 5193 population controls belonging to the National Blood Donor Service (NBS) and the 1958 Birth Cohort (BC58) and 929 cancer-free controls from our COloRectal Gene Identification study (CORGI). The NBS and BC58 samples had been genotyped with Illumina 1.2M (Hap1.2M) arrays as part of the Wellcome Trust Case Control Consortium 2, and the CORGI controls had been genotyped with Illumina Hap550 arrays (Hap550, N=932) as part of our ongoing studies in colorectal cancer genetics. Two of the five SNPs (rs2910164 and rs1867277) were not included on the Illumina 1.2M and Hap550 arrays, but had excellent proxy markers that facilitated their imputation. rs2910164 is perfectly tagged (D'=1, r²=1) by rs2961920, a marker present on both 9q22 and 14q13 and risk allele SNP, genotypes and risk allele

<table>
<thead>
<tr>
<th>SNP, genotypes and risk allele</th>
<th>Frequency (%)</th>
<th>ORs for genotype or per allele overall (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2910164</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>436 (0.578)</td>
<td>3540 (0.584)</td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>271 (0.367)</td>
<td>2178 (0.360)</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>41 (0.054)</td>
<td>339 (0.056)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (C)</td>
<td>359 (0.238)</td>
<td>2657 (0.238)</td>
<td></td>
</tr>
<tr>
<td>rs6985267</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>164 (0.218)</td>
<td>1441 (0.236)</td>
<td></td>
</tr>
<tr>
<td>GT</td>
<td>346 (0.461)</td>
<td>3012 (0.493)</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>241 (0.321)</td>
<td>1662 (0.272)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (G)</td>
<td>674 (0.449)</td>
<td>5894 (0.518)</td>
<td></td>
</tr>
<tr>
<td>rs965513</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>187 (0.249)</td>
<td>2748 (0.449)</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>394 (0.525)</td>
<td>2729 (0.446)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>170 (0.226)</td>
<td>642 (0.105)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (A)</td>
<td>734 (0.489)</td>
<td>4015 (0.328)</td>
<td></td>
</tr>
<tr>
<td>rs1867277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>159 (0.211)</td>
<td>2260 (0.376)</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>396 (0.529)</td>
<td>2879 (0.473)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>196 (0.260)</td>
<td>918 (0.151)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (A)</td>
<td>790 (0.525)</td>
<td>4715 (0.387)</td>
<td></td>
</tr>
<tr>
<td>rs944289</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>87 (0.116)</td>
<td>1003 (0.164)</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>332 (0.441)</td>
<td>2924 (0.478)</td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>334 (0.444)</td>
<td>2193 (0.358)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (T)</td>
<td>1000 (0.664)</td>
<td>7310 (0.597)</td>
<td></td>
</tr>
</tbody>
</table>

**Quality control and statistical analysis**

General genotyping quality control assessment was as previously described. Duplicate samples were used to check genotyping quality and 100% concordance was found. Samples with multiple missing genotypes were eliminated from the analyses (N=14). All five SNPs passed our quality control thresholds including call rates >95% and Hardy-Weinberg equilibrium p values >0.05.

Association statistics were obtained on per allele, genotypic and haplotype bases using logistic regression models implemented in PLINK, R, and SNPTREE. Haplotype analyses were carried out with HAPLOVIEW and PLINK. Allelic count association meta-analyses, using the Mantel-Haenszel method, were carried out in STATA. We used the IMPUTE2 software formally to generate rs2910164 and rs1867277 genotypes in the control population, although the fact that perfect proxies were used rendered this task of very limited utility.

**RESULTS**

**Single SNP analyses**

Three of the five SNPs examined showed a significant association with TC risk in the UK population (table 1). The strongest associations, on a per allele basis, were observed for the 9q22 SNPs rs695513A (p=6.55×10⁻³⁴, OR=1.99, 95% CI 1.77 to 2.21) and rs1867277A (p=5.90×10⁻²⁴, OR=1.75, 95% CI 1.57 to 1.95). The association at rs944289T on 14q15 was also convincingly replicated (p=6.95×10⁻⁷, OR=1.53, 95% CI 1.18 to 1.48). For rs6985267G, we also found a nominally significant

<p>| Table 1 Association statistics for thyroid cancer risk and genetic variants at chromosomes 5q24, 8q24, 9q22 and 14q13 |
|--------------------------------------------------|---------------------------------------------|--------|</p>
<table>
<thead>
<tr>
<th>SNP, genotypes and risk allele</th>
<th>Frequency (%)</th>
<th>ORs for genotype or per allele overall (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2910164</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>436 (0.578)</td>
<td>3540 (0.584)</td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>271 (0.367)</td>
<td>2178 (0.360)</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>41 (0.054)</td>
<td>339 (0.056)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (C)</td>
<td>359 (0.238)</td>
<td>2657 (0.238)</td>
<td></td>
</tr>
<tr>
<td>rs6985267</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>164 (0.218)</td>
<td>1441 (0.236)</td>
<td></td>
</tr>
<tr>
<td>GT</td>
<td>346 (0.461)</td>
<td>3012 (0.493)</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>241 (0.321)</td>
<td>1662 (0.272)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (G)</td>
<td>674 (0.449)</td>
<td>5894 (0.518)</td>
<td></td>
</tr>
<tr>
<td>rs965513</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>187 (0.249)</td>
<td>2748 (0.449)</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>394 (0.525)</td>
<td>2729 (0.446)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>170 (0.226)</td>
<td>642 (0.105)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (A)</td>
<td>734 (0.489)</td>
<td>4015 (0.328)</td>
<td></td>
</tr>
<tr>
<td>rs1867277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>159 (0.211)</td>
<td>2260 (0.376)</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>396 (0.529)</td>
<td>2879 (0.473)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>196 (0.260)</td>
<td>918 (0.151)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (A)</td>
<td>790 (0.525)</td>
<td>4715 (0.387)</td>
<td></td>
</tr>
<tr>
<td>rs944289</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>87 (0.116)</td>
<td>1003 (0.164)</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>332 (0.441)</td>
<td>2924 (0.478)</td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>334 (0.444)</td>
<td>2193 (0.358)</td>
<td></td>
</tr>
<tr>
<td>Risk allele (T)</td>
<td>1000 (0.664)</td>
<td>7310 (0.597)</td>
<td></td>
</tr>
</tbody>
</table>

SNP, single nucleotide polymorphism.
Cancer genetics

Table 2  Thyroid cancer risk associated with different haplotypes, defined by rs965513 and rs1867277 alleles, at chromosome 9q22

<table>
<thead>
<tr>
<th>Haplotype rs965513</th>
<th>rs1867277</th>
<th>Frequency</th>
<th>OR (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G</td>
<td>G</td>
<td>0.413</td>
<td>0.561</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>A</td>
<td>0.428</td>
<td>0.276</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>G</td>
<td>0.097</td>
<td>0.111</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>A</td>
<td>0.061</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Haplotype frequencies were estimated using Haploviz (http://www.haploviz.org). Only samples with full data at both loci were used for the analyses (761 cases and 6065 controls).

association with TC risk (p=0.016, OR=1.14, 95% CI 1.02 to 1.27, equivalent false discovery rate=0.020). However, the association between TC and rs2910164 was not replicated (P_{allelic}=0.85, OR=1.01, 95% CI 0.89 to 1.14). To test if the strength of these associations were similar in cases with different histological types, we carried out cases-only interaction analyses. We found no differences between cases with papillary and follicular histology (p=0.25 for all markers, data not shown), suggesting that these associations were not restricted to any particular histological type of TC.

Are there multiple risk alleles on chromosome 9q22?
The 9q22 variants associated with TC risk, rs965513 and rs1867277, were originally reported by independent and non-overlapping GWA and candidate gene studies. Other studies have not considered whether these SNPs represent independent signals of association. We performed unconditional logistic regression analyses incorporating rs1867277 and rs965513 genotypes as variables and sex as a covariate in the model. Both the rs965513 (p=2.34×10^{-13}, OR=1.74) and rs1867277 (p=0.008, OR=1.21) association signals decreased but remained significant. We then reconstructed haplotypes at these two loci and estimated the ORs associated with having each one of the three possible risk haplotypes (haplotype 2 = rs965513A-rs1867277A, haplotype 3 = rs965513G-rs1867277A, and haplotype 4=rs965513A-rs1867277G, table 2) compared with the non-risk haplotype (haplotype 1 = rs965513G-rs1867277G). As expected, carrying the haplotype with both risk alleles (haplotype 2) increased disease risk significantly (p=2.19×10^{-35}, OR=2.09). Carrying haplotypes with either one risk allele at rs965513 (haplotype 3) or at rs1867277 (haplotype 4) also increased risk, although the association signal for the haplotype 3 was weaker (P_{haplotype 3}=0.07, OR_{haplotype 3}=1.19 and P_{haplotype 4}=0.0001, OR_{haplotype 4}=1.61, table 2).

We also estimated the risk associated with ‘diplotypes’ at each the two 9q22 loci. Table 3 shows the genotype frequencies at the two SNPs and the ORs associated with the nine possible diplotypes. Individuals with the four risk alleles at both loci (~7.4% of the general population) had a 4.45-fold higher risk than non-carriers (~31.5% of the population), with the other diplotypes having intermediate risk levels, principally dependent on rs965513 (table 3).

These analyses showed that the two SNPs did not simply and efficiently tag a single high-risk haplotype on 9q22. However, they did not distinguish between the existence of multiple independent risk alleles at 9q22 and a third ‘causal’ variant tagged in complex fashion by both rs1867277 and rs965513. To undertake a limited examination of the latter possibility, we searched for SNPs in high/moderate LD (r²>0.5) with both rs1867277 and rs965513 in the most recent release of the 1000 Genomes Project (phase 1, interim release, 11 May 2011, n=762 European samples). We identified four such SNPs (rs10124220, rs7849873, rs6474813, and rs1444352, supplementary table 2, supplementary figure 1). However, none of these polymorphisms lay at a site with evidence of functional importance (data not shown). We found no evidence for a role of non-synonymous variants within any of the seven nearby 9q22 genes (supplementary figure 1).

TC is associated with variation at 8q24 under a recessive model
We found evidence that rs6983267G was associated with TC in the UK population (table 1). Interestingly, and unlike previous findings in other cancer types, rs6983267G was associated with TC risk according to a recessive model (tables 1 and 4). We found no difference in risk between non-carriers and heterozygotes (OR=1.01, 95% CI 0.83 to 1.24, p=0.921, table 1), but a significantly increased risk when homozygous carriers (GG) were compared to non-carriers (p=0.016, OR=1.14, table 1), to heterozygotes (p=0.009, OR=1.27, table 1), and to non-carriers/ heterozygotes (p=0.004, OR=1.26, table 4). Wokolorczyk et al had previously found relatively weak evidence of association between TC and the rs6983267 SNP in the Polish population.

Table 3  Thyroid cancer risk associated with different genotype combinations (diplotypes) at rs965513 and rs1867277

<table>
<thead>
<tr>
<th>rs965513</th>
<th>rs1867277</th>
<th>Frequency (%)</th>
<th>OR (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GG</td>
<td>GG</td>
<td>123 (0.162)</td>
<td>1917 (0.315)</td>
<td>Reference</td>
</tr>
<tr>
<td>AG</td>
<td>AA</td>
<td>60 (0.079)</td>
<td>744 (0.122)</td>
<td>1.257 (0.897 to 1.747)</td>
</tr>
<tr>
<td>AA</td>
<td>GG</td>
<td>7 (0.009)</td>
<td>73 (0.012)</td>
<td>1.494 (0.568 to 3.333)</td>
</tr>
<tr>
<td>AG</td>
<td>GG</td>
<td>33 (0.043)</td>
<td>355 (0.058)</td>
<td>1.449 (0.939 to 2.183)</td>
</tr>
<tr>
<td>AG</td>
<td>AA</td>
<td>300 (0.394)</td>
<td>1960 (0.322)</td>
<td>2.385 (1.908 to 2.995)</td>
</tr>
<tr>
<td>AA</td>
<td>GG</td>
<td>63 (0.083)</td>
<td>394 (0.065)</td>
<td>2.491 (1.774 to 3.473)</td>
</tr>
<tr>
<td>AA</td>
<td>AG</td>
<td>4 (0.005)</td>
<td>17 (0.003)</td>
<td>3.663 (0.883 to 11.465)</td>
</tr>
<tr>
<td>AA</td>
<td>AA</td>
<td>42 (0.055)</td>
<td>174 (0.029)</td>
<td>3.759 (2.497 to 5.580)</td>
</tr>
<tr>
<td>AA</td>
<td>AA</td>
<td>129 (0.170)</td>
<td>451 (0.074)</td>
<td>4.455 (3.379 to 5.876)</td>
</tr>
</tbody>
</table>

The risk alleles are rs965513G and rs1867277A.
rs2910164 at the pre-miR-146a locus is not associated with TC
Using an allelic model, our study found no evidence of an
association between rs2910164 and TC risk (p=0.85, table 1). We
confirmed the absence of associations between this SNP and
TC in genotypic, dominant, recessive and trend models (p>0.71
for all models, supplementary table 5). The previous report of an
association between rs2910164 and papillary TC risk is not
supported by our study (p=0.784, supplementary table 3). When the analyses were restricted to the cases that had histologically verified papillary TC we also failed to detect association between rs2910164 and TC (p=0.84
for all models, data not shown, supplementary table 5).

Combined effects of rs6983267, rs965513, rs1867277, and rs944289 on disease risk
We carried out case-only and case-control pairwise analyses
between all four risk SNPs associated with TC and found to no
evidence for SNP–SNP interaction (details not shown). We then
estimated the combined effects of the four SNPs on risk. To
incorporate the effects of the two 9q22 markers in the combined
risk analyses, we used the estimates obtained in our diplotype
analyses (table 3). Using this information, we estimated that the
risk for those individuals who are homozygous at 8q24, 9q22 and
14q13—comprising ~1.7% of the UK population—is 9.96-fold
higher compared with individuals who do not carry any risk
allele at these loci (~1.2% of the population, table 5). The risk
homozygous at the four SNPs (17 cases and 41 controls,
supplementary table 5) have a 17.08-fold higher chance of having
homozygous at the four SNPs (17 cases and 41 controls,
supplementary table 5) have a 17.08-fold higher chance of having
risk between heterozygous and non-carriers (OR=1.087, p=0.142, table 4).

rs2910164 at the pre-miR-146a locus is not associated with TC
Using an allelic model, our study found no evidence of an
association between rs2910164 and TC risk (p=0.85, table 1). We
confirmed the absence of associations between this SNP and
TC in genotypic, dominant, recessive and trend models (p>0.71
for all models, supplementary table 5). The previous report of an
association between rs2910164 and papillary TC risk is not
supported by our study (p=0.784, supplementary table 3). When the analyses were restricted to the cases that had histologically verified papillary TC we also failed to detect association between rs2910164 and TC (p=0.84
for all models, data not shown, supplementary table 5).

Combined effects of rs6983267, rs965513, rs1867277, and rs944289 on disease risk
We carried out case-only and case-control pairwise analyses
between all four risk SNPs associated with TC and found to no
evidence for SNP–SNP interaction (details not shown). We then
estimated the combined effects of the four SNPs on risk. To
incorporate the effects of the two 9q22 markers in the combined
risk analyses, we used the estimates obtained in our diplotype
analyses (table 3). Using this information, we estimated that the
risk for those individuals who are homozygous at 8q24, 9q22 and
14q13—comprising ~1.7% of the UK population—is 9.96-fold
higher compared with individuals who do not carry any risk
allele at these loci (~1.2% of the population, table 5). The risk
homozygous at the four SNPs (17 cases and 41 controls,
supplementary table 5) have a 17.08-fold higher chance of having
homozygous at the four SNPs (17 cases and 41 controls,
supplementary table 5) have a 17.08-fold higher chance of having
(p=0.04, OR=0.78, 95% CI 0.63 to 1.00). We carried out a meta-
analysis of the Polish data and our data and found enhanced
support for an association between rs6983267G and TC risk,
with no evidence of inter-study heterogeneity (Pmeta=6.64×10^{-4}, per allele ORmeta=1.15, 95% CI 1.06 to 1.25, P_heterogeneity=0.341, supplementary figure 2). The meta-analysis
continued to support a recessive effect of the rs6983267G allele
on risk (OR >1.2, p<0.004, table 4) and found no difference in
risk between heterozygous and non-carriers (OR=1.087, p=0.142, table 4).

Table 4 Evidence that the association between rs6983267 and thyroid cancer risk is best explained by a
recessive model

<table>
<thead>
<tr>
<th>Test</th>
<th>UK only OR (95% CI)</th>
<th>p Value</th>
<th>Meta analysis of UK and Polish studies OR (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GG vs (GT + TT)</td>
<td>1.286 (1.071 to 1.494)</td>
<td>0.004</td>
<td>1.250 (1.089 to 1.435)</td>
<td>7.64×10^{-4}</td>
</tr>
<tr>
<td>GG vs GT</td>
<td>1.282 (1.055 to 1.509)</td>
<td>0.009</td>
<td>1.215 (1.051 to 1.404)</td>
<td>0.004</td>
</tr>
<tr>
<td>GT vs TT</td>
<td>1.010 (0.827 to 1.236)</td>
<td>0.960</td>
<td>1.087 (0.933 to 1.266)</td>
<td>0.142</td>
</tr>
</tbody>
</table>

Note that the risk allele homozygotes (GG) have significantly higher frequency compared with other genotypes combined (GT + TT) and with heterozygotes (GT). This is also the case for the GG versus TT test which is not shown. However, there is no evidence that heterozygotes are over-represented in cases (OR vs TT).

10.9% of the sibling relative risk of TC. Under a conser-

vative model that assumes the existence of a single risk variant
at chromosome 9q22, these loci explain at least 6.6% of the
disease heritability (supplementary table 4).

DISCUSSION
Using a relatively large sample set in a single, homogeneous
European population, we confirmed associations between TC
and SNPs on chromosomes 9q22 (rs965513 and rs1867277),
14q13 (rs944289), and 8q24 (rs6983267). However, we failed to
replicate an association between SNP rs2910164 on 8q24 and
TC risk.

The 9q22 SNPs rs965513 and rs1867277 have not previously
been genotyped in the same samples. We have found that there
is not a single TC risk haplotype on 9q22 that is perfectly
denoted by rs965513 and rs1867277. Conversely, the association
cannot be explained entirely by genotypes at only one of the
two SNPs, although logistic regression analysis incorporating
both SNPs did lead to a considerably reduced association signal
for both SNPs, particularly rs1867277. We suggest, therefore,
that rs965513 and rs1867277 tag a third variant (or variants) that
is the functional variation near FOXE1. Perhaps contrary to the
gene, Landa et al. showed that rs1867277A affected
FOXE1 transcript levels through the differential recruitment of
the USF1/USF2 transcription factors and suggested that
rs1867277 was a TC-causal SNP. Fine mapping studies at this
site might benefit from the use of non-European samples. For
example, the LD between rs965513 and rs1867277 is signifi-
cantly weaker in populations of African (r^2=0.01, D'=0.078) or
Asian (r^2=0.00, D'=0.014) ancestry (data from the 1000 Genomes
Project). Finally, although challenging, rare variants at
FOXE1 deserve further scrutiny.

Our study validated the association between TC and
rs6983267 and extended the range of cancer types associated
with this variant. Interestingly, however, we found that
rs6983267G is associated with TC risk in a recessive fashion; all
other rs6983267 cancer associations follow an allelic dosage
model. Recessive cancer predisposition SNPs have rarely been

Table 5 Estimates to genotype relative risk at rs6983267, 9q22, and
rs944289

<table>
<thead>
<tr>
<th>Locus</th>
<th>Non-carriers Population frequency</th>
<th>Heterozygous Population frequency</th>
<th>Homozygous carriers Population frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6983267*</td>
<td>0.236</td>
<td>NA</td>
<td>1.274</td>
</tr>
<tr>
<td>rs922†</td>
<td>0.315</td>
<td>2.385</td>
<td>4.455</td>
</tr>
<tr>
<td>rs944289</td>
<td>0.104</td>
<td>1.309</td>
<td>1.755</td>
</tr>
<tr>
<td>Combined</td>
<td>0.012</td>
<td>3.122</td>
<td>9.961</td>
</tr>
</tbody>
</table>

*rs6983267 heterozygous do not have increased risk of thyroid cancer (see tables 1 and 4).
† ORs and frequencies for the 9q22 markers were obtained from the diplotype analysis presented in table 3. Population frequency from non-carriers, heterozygous and homozygous carriers is shown in table 3.
Cancer genetics

found. In part, this probably reflects suboptimal power in GWA studies and our finding emphasises the role of candidate SNP analyses across cancer types. Given that rs6983267 itself may be functional in predisposition to colorectal and other cancers,\textsuperscript{17} one possibility is that the true, recessive functional variation in TC is not rs6983267, but an SNP in strong LD with it.

The first association between TC and common genetic variants was found at the pre-mir-146a locus (4), a micro-RNA that is upregulated in thyroid tumours.\textsuperscript{18} There was no good evidence in our study of an association between rs2910164 and TC risk. One other possibility is that deviations from Hardy-Weinberg equilibrium were present in the case genotypes of Jazdewski et al.; it is not clear whether this was the result or the cause of the heterozygote-association with TC risk. Other possible explanations for the differences between Jazdewski et al’s study and our own include chance, systematic differences between cases and controls (whether related to ascertainment or technical issues) and population specific effects in either study.

The four validated TC risk SNPs explain an approximately 10-fold differential risk between those with high risk alleles and those with all low risk alleles. Moreover, owing to the large effect size associated with the 9q22 SNPs, the four SNPs explain over 10% of the total sibling relative risk of TC, despite the fact that TC is one of the largest of familial relative risks reported for any malignancy. It is highly plausible that future studies involving only a few thousand cases and controls could identify additional important common risk variants for this common disease.

Author affiliations

1. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
2. Department of Endocrine Surgery, John Radcliffe Hospital, Oxford, UK
3. Velindre Cancer Centre, Cardiff, UK
4. Mount Vernon Hospital, Northwood, UK
5. Institute of Head and Neck Studies and Education, University Hospitals of Coventry and Warwickshire, Walsgrave, Coventry, UK
6. Guy’s and St Thomas’ NHS Foundation Trust and King’s College London, London, UK
7. Weston Park Hospital, Sheffield, UK
8. Kent and Canterbury Hospital, Canterbury, UK
9. York Hospital, York, UK
10. Bristol Hematology and Oncology Centre, Bristol, UK
11. Norfolk and Norwich University NHS Trust, Norwich, UK
12. Medway Maritime Hospital, Gillingham, UK
13. St James University Hospital, Leeds, UK
14. St Mary’s Hospital, London, UK
15. Newcross Hospital, Wolverhampton, UK
16. NHRI Comprehensive Biomedical Research Centre, University of Oxford, Oxford, UK

Acknowledgements

We are grateful to all of the individuals who participated in the TCUKIN study. We are also grateful to the National Cancer Research Network and to the National Cancer Research Institute’s Thyroid Cancer Subgroup for supporting the TCUKIN study. We acknowledge the help of the Wellcome Trust Case-Control Consortium in making control data publicly available.

Funding

Cancer Research UK provided principal funding for this study. LGC and IT receive funding from the FP7 CHIBCHA Consortium. The Wellcome Trust Centre for Human Genetics is funded by the Wellcome Trust (Grant number: 075491/Z/04).

Correction notice

This paper has been corrected since it was first published online. The corresponding author’s name should read Luis G Carvajal-Carmona.

Competing interests

None.

Patient consent

Obtained.

Ethics approval

Ethics approval was provided by the Southampton and South West Hampshire Research Ethics Committee (A).

Contributors

LGC and IT are the leaders of TCUKIN and designed the study, performed the main statistical analyses and co-wrote the manuscript. AMJ, KMH, LM and MG carried out the experiments and coordinated sample collection. AA carried out some of the statistical analyses. CH, HM, MH, SMC, JW, EM, AC, MB, TR, CM, PR, GG, DP, and CB are members of the TCUKIN Consortium, coordinated sample collection, and summarised the clinical data. All authors contributed to the final version of the manuscript.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data sharing statement

The data presented in the manuscript are available online.

REFERENCES


12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker P, Daly MJ, Shimo PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.


APPENDIX 1
Collaborators in the TCUKIN Study include: Dr Laura Moss, Velindre Cancer Centre, Cardiff CF14 2TL, UK; Dr Christopher Scrase, The Ipswich Hospital, Ipswich, IP4 6PD, UK; Dr Andrew Goodman, Royal Devon & Exeter Hospital, Exeter, EX2 5DQ, UK; Dr Radu Mihai, John Radcliffe Hospital, Oxford OX3 9DU, UK; Dr James Gildersleeve, Royal Berkshire Hospital, Reading RG1 5AN, UK; Dr Catherine Lemon, Mount Vernon Hospital, Northwood, HA6 2RN, UK; Dr Antony Robinson, Royal United Hospital, Bath BA1 3NG, UK; Dr Caroline Brammer, Newcross Hospital, Wolverhampton WV10 0QP, UK; Dr Georgina Gerrard, St. James University Hospital, Leeds LS9 7TF, UK; Professor Hisham Mehmanna, Institute of Head and Neck Studies and Education, University Hospitals of Coventry and Warwickshire, Walsgrave, Coventry CV2 2DX, UK; Dr Matthew Beasley, Bristol Hematology and Oncology Centre, Bristol BS2 8ED, UK; Dr Hosahalli K Mohan and Dr Susan EM Clarke, Guys and St Thomas’ NHS Foundation Trust and King’s College London, London SE1 9RT, UK; Dr Kate Goodchild, Luton & Dunstable Hospital, Luton LU4 0DZ, UK; Dr Jonathan Wadley, Weston Park Hospital, Sheffield S10 2SZ, UK; Dr Abdel Hamid, Stunthorpe General Hospital, Stunthorpe DN15 7BH, UK; Dr Danielle Power, St. Mary’s Hospital, London W2 1NY, UK; Dr Elena Macias, Kent and Canterbury Hospital, Canterbury, CT1 3NG, UK; Dr Jerry Sharp, Royal Derby Hospital, Derby DE22 3NE, UK; Dr Andrew Coatsworth, York Hospital, York Y031 8HW, UK; Dr Hamish Courtney, Royal Victoria Hospital, Belfast BT12 6BA, UK; Dr Stephen Whitaker and Dr Katie Wood, Royal Surrey County Hospital, Guildford GU1 7XX, UK; Dr James McCaull, Bradford Royal Infirmary, Bradford BD9 6RJ, UK; Dr Christopher Ashford, Worcestershire Royal Hospital, Worcester WR5 1DD, UK; Dr Tom Roques and Dr Craig Martin, Norfolk and Norwich University Hospital NHS Trust, Norwich NR4 7UY, UK; Dr Vivienne Loo, Broomfield Hospital, Chelmsford CM1 7ET, UK; Dr Jennifer Marshall, Southampon General Hospital, Southampton S016 6YD, UK; Dr Amy Roy, Derriford Hospital, Plymouth PL6 8DH, UK; Dr Joanna Simpson, The Royal Sussex County Hospital, Brighton BN2 5BE, UK; Dr Nick Rowell, Maidstone Royal Hospital, Maidstone ME16 9QQ, UK; Mr Simon Ellenbogen, Tameside General Hospital, Ashton-under-Lyne OL6 0RW, UK; Dr Paul Ryan, Medway Maritime Hospital, Gillingham ME7 5NY, UK; Dr Arshad Jamil, University Hospital North Staffs, Stoke on Trent ST4 6GQ, UK.

Journal of Medical Genetics impact factor

Journal of Medical Genetics is delighted to announce that its latest impact factor has been published in the 2010 Journal Citation report (Thomson Reuters, 2011) as 7.037 — an increase from 5.751 last year, which reflects the high quality of the publication in the genetics field. If you would like to add to this quality, then submit your paper at jmg.bmj.com.