Risk of non-medullary thyroid cancer influenced by polymorphic variation in the thyroglobulin gene

Athena Matakidou1,2, Nancy Hamel3, Sanjay Popat4, Kiersten Henderson5, Tania Kantemiroff5, Clive Harmer5, Susan E.M.Clarke6, Richard S.Houlston1,8 and William D.Foulkes3,4,7

1Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK, 2Department of Medicine, Royal Marsden Hospital NHS Trust, SM2 5TT, UK, 3Cancer Prevention Center, Sir MB Davis—Jewish General Hospital, McGill University, Montreal, Quebec, H3T 1E2, Canada, 4Division of Medical Genetics, Department of Medicine, Research Institute of the McGill University Montreal General Hospital Research Institute, Montreal, Quebec, H3G 1A2, Canada, 5Department of Clinical Oncology, Royal Marsden Hospital NHS Trust, SM2 5NG, UK, 6Department of Nuclear Medicine, Guy’s and St Thomas’ Hospital, London SE1 9RT, UK and 7Program in Cancer Genetics, Departments of Oncology and Human Genetics, McGill University, Montreal, Quebec, H2W 1S6, Canada

8To whom correspondence should be addressed
Email: richard.houlston@icr.ac.uk

Benign thyroid disorders are strong risk factors for non-medullary thyroid cancer (NMTC). Germline variation in Tg (thyroglobulin) and TSHR (thyroid stimulating hormone receptor) confers an increased risk of benign thyroid disorders. To explore the hypothesis that polymorphic variation in these genes affects the risk of NMTC we compared the frequency of TgQ2511R, TSHR-P52T and TSHR-D727E genotypes in two series of NMTC cases and controls (group 1, Canadian 102 cases and 102 controls; group 2, British 202 cases and 298 controls). No significant association was seen with TSHR-P52T and TSHR-D727E genotypes and risk of NMTC. However, the frequency of the R-allele of TgQ2511R was over represented in NMTC cases (~5-8-fold in first-degree relatives of NMTC cases (3)). Four familial NMTC loci have been identified through linkage analysis of single families with multiple affected family members. The TCO locus (19p13.2) was identified in a French pedigree with multiple cases of multinodular goiter (MNG) and NMTC (4). The fPTC locus was mapped to 1q21 in a large, three-generation pedigree in which both papillary thyroid cancer (PTC) and renal papillary carcinoma segregated (5). MNG1 (14q32) was identified through the study of a single, large Canadian pedigree with 18 MNG cases and two PTC cases (6). A fourth locus has recently been localized on chromosome 2q21 in a large Tasmanian family (7). These linkages have been detected in large single families and other families have not shown linkage to these loci (6,8,9), suggesting that the disease is genetically highly heterogeneous. It is therefore very unlikely that sequence changes at these loci will account for most of the familial risk seen in relatives of cases. Whilst part of the residual familial risk could be due to high penetrance mutations in as yet unidentified genes, a polygenic mechanism may provide a more plausible alternative explanation. Alleles conferring relative risks of ~2 or less will rarely cause multiple-case families and are difficult or impossible to identify through linkage (10). The search for low penetrance alleles has therefore centred on association studies based on comparing the frequency of polymorphic genotypes in cases and controls. Alleles positively associated with the disease are analogous to risk factors in epidemiology and may be causally related to disease risk or in linkage disequilibrium with disease-causing variants (11).

Thyroglobulin (Tg) is the most highly expressed protein in the thyroid gland and functions as a scaffold protein for thyroid hormonogenesis and as storage protein for thyroid hormones and iodide. Tg mRNA levels are low in NMTC (12), and germline Tg abnormalities have been reported in association with goiter (13–15). Goiter is frequently observed in NMTC pedigrees and is an established risk factor for NMTC in general (16).

TSHR (thyroid stimulating hormone receptor) is a G-protein-coupled receptor and binds the thyroid stimulating hormone (TSH), the main stimulator of growth and function of normal follicular cells. Germline mutations in TSHR have been shown to co-segregate with hyperthyroidism and thyroid hyperplasia in six families (17). Mutated forms of TSHR and adenylyl cyclase-activating Gs protein have also been observed in thyroid tumors (18–21).

Given the role of the Tg and TSHR genes in non-malignant thyroid disease, these represent strong candidate NMTC susceptibility loci.

It is conceivable that subtle variations in Tg or TSHR activity, a consequence of polymorphic variation, may confer an increased risk of NMTC. The sequence change 7589A→G in Tg is the basis of the Q2511R polymorphism (22). Two polymorphisms in TSHR have been associated previously with studies have shown that the risk of NMTC is increased ~5-8-fold in first-degree relatives of NMTC cases (3). Four familial NMTC loci have been identified through linkage analysis of single families with multiple affected family members. The TCO locus (19p13.2) was identified in a French pedigree with multiple cases of multinodular goiter (MNG) and NMTC (4). The fPTC locus was mapped to 1q21 in a large, three-generation pedigree in which both papillary thyroid cancer (PTC) and renal papillary carcinoma segregated (5). MNG1 (14q32) was identified through the study of a single, large Canadian pedigree with 18 MNG cases and two PTC cases (6). A fourth locus has recently been localized on chromosome 2q21 in a large Tasmanian family (7). These linkages have been detected in large single families and other families have not shown linkage to these loci (6,8,9), suggesting that the disease is genetically highly heterogeneous. It is therefore very unlikely that sequence changes at these loci will account for most of the familial risk seen in relatives of cases. Whilst part of the residual familial risk could be due to high penetrance mutations in as yet unidentified genes, a polygenic mechanism may provide a more plausible alternative explanation. Alleles conferring relative risks of ~2 or less will rarely cause multiple-case families and are difficult or impossible to identify through linkage (10). The search for low penetrance alleles has therefore centred on association studies based on comparing the frequency of polymorphic genotypes in cases and controls. Alleles positively associated with the disease are analogous to risk factors in epidemiology and may be causally related to disease risk or in linkage disequilibrium with disease-causing variants (11).

Thyroglobulin (Tg) is the most highly expressed protein in the thyroid gland and functions as a scaffold protein for thyroid hormonogenesis and as storage protein for thyroid hormones and iodide. Tg mRNA levels are low in NMTC (12), and germline Tg abnormalities have been reported in association with goiter (13–15). Goiter is frequently observed in NMTC pedigrees and is an established risk factor for NMTC in general (16).

TSHR (thyroid stimulating hormone receptor) is a G-protein-coupled receptor and binds the thyroid stimulating hormone (TSH), the main stimulator of growth and function of normal follicular cells. Germline mutations in TSHR have been shown to co-segregate with hyperthyroidism and thyroid hyperplasia in six families (17). Mutated forms of TSHR and adenylyl cyclase-activating Gs protein have also been observed in thyroid tumors (18–21).

Given the role of the Tg and TSHR genes in non-malignant thyroid disease, these represent strong candidate NMTC susceptibility loci.

It is conceivable that subtle variations in Tg or TSHR activity, a consequence of polymorphic variation, may confer an increased risk of NMTC. The sequence change 7589A→G in Tg is the basis of the Q2511R polymorphism (22). Two polymorphisms in TSHR have been associated previously with studies have shown that the risk of NMTC is increased ~5-8-fold in first-degree relatives of NMTC cases (3). Four familial NMTC loci have been identified through linkage analysis of single families with multiple affected family members. The TCO locus (19p13.2) was identified in a French pedigree with multiple cases of multinodular goiter (MNG) and NMTC (4). The fPTC locus was mapped to 1q21 in a large, three-generation pedigree in which both papillary thyroid cancer (PTC) and renal papillary carcinoma segregated (5). MNG1 (14q32) was identified through the study of a single, large Canadian pedigree with 18 MNG cases and two PTC cases (6). A fourth locus has recently been localized on chromosome 2q21 in a large Tasmanian family (7). These linkages have been detected in large single families and other families have not shown linkage to these loci (6,8,9), suggesting that the disease is genetically highly heterogeneous. It is therefore very unlikely that sequence changes at these loci will account for most of the familial risk seen in relatives of cases. Whilst part of the residual familial risk could be due to high penetrance mutations in as yet unidentified genes, a polygenic mechanism may provide a more plausible alternative explanation. Alleles conferring relative risks of ~2 or less will rarely cause multiple-case families and are difficult or impossible to identify through linkage (10). The search for low penetrance alleles has therefore centred on association studies based on comparing the frequency of polymorphic genotypes in cases and controls. Alleles positively associated with the disease are analogous to risk factors in epidemiology and may be causally related to disease risk or in linkage disequilibrium with disease-causing variants (11).

Thyroglobulin (Tg) is the most highly expressed protein in the thyroid gland and functions as a scaffold protein for thyroid hormonogenesis and as storage protein for thyroid hormones and iodide. Tg mRNA levels are low in NMTC (12), and germline Tg abnormalities have been reported in association with goiter (13–15). Goiter is frequently observed in NMTC pedigrees and is an established risk factor for NMTC in general (16).

TSHR (thyroid stimulating hormone receptor) is a G-protein-coupled receptor and binds the thyroid stimulating hormone (TSH), the main stimulator of growth and function of normal follicular cells. Germline mutations in TSHR have been shown to co-segregate with hyperthyroidism and thyroid hyperplasia in six families (17). Mutated forms of TSHR and adenylyl cyclase-activating Gs protein have also been observed in thyroid tumors (18–21).

Given the role of the Tg and TSHR genes in non-malignant thyroid disease, these represent strong candidate NMTC susceptibility loci.

It is conceivable that subtle variations in Tg or TSHR activity, a consequence of polymorphic variation, may confer an increased risk of NMTC. The sequence change 7589A→G in Tg is the basis of the Q2511R polymorphism (22). Two polymorphisms in TSHR have been associated previously with
non-malignant thyroid disease—the P52T and autoimmune thyroid disease (23) and the D727E with goiter (24).

To examine the proposition that genetic variation in TG or TSHR, defined by these polymorphisms, is a determinant of NMTC risk we compared the frequency of these in cases and controls from two study populations.

Materials and methods

Subjects

Two study groups were analyzed. Table I details the demographic features of the patients and controls from both the study populations. In both groups cases had historically proven NMTC—papillary, follicular, Hurtle cell, or mixed papillary and follicular cancers. The study was undertaken with local ethical board approval at each center in accordance with the tenets of the Helsinki accord.

Study group 1: Canadian cohort. A total of 249 prevalent cases of NMTC diagnosed at ENT clinics of the Sir Mortimer B Davis–Jewish General Hospital and the Montreal General Hospital, Montreal between 1986 and 1996 were eligible for the study. This represents around one-third of all NMTC cases diagnosed in the Montreal area of Canada during this period. Of these, 71 were not included in the study for the following reasons—26 refused to participate, 43 were untraceable, two were deemed unable to complete a questionnaire and provide informed consent. Of the 178 remaining 153 provided a blood sample. Cases who could not be contacted or did not participate were not significantly different from the rest of the cohort in terms of age, gender or histology.

A series of outpatient controls—patients (with non-thyroid related disease) and spouses—were recruited. Exclusion criteria included a history of benign or malignant thyroid disease or malignancy of any type (except non-melanotic skin cancer). Each control was matched to a case for sex, year of birth (±5 years) and ancestral country of origin (at least two of four grandparents of controls had to have the same ethnic origin as the case). All were resident within the Montreal area.

Twenty-nine samples had incomplete data for at least one of three polymorphic loci after molecular analysis, and one case and nine controls did not provide informed consent. Of the 178 remaining 153 provided a blood sample. 43 were untraceable, two were deemed unable to complete a questionnaire and provide informed consent. Of the 178 remaining 153 provided a blood sample. Cases who could not be contacted or did not participate were not significantly different from the rest of the cohort in terms of age, gender or histology.

Two study groups were analyzed. Table I details the demographic features of the patients and controls from both the study populations. In both groups cases had historically proven NMTC—papillary, follicular, Hurtle cell, or mixed papillary and follicular cancers. The study was undertaken with local ethical board approval at each center in accordance with the tenets of the Helsinki accord.

Study group 2: British Isles cohort. 202 patients (incident and prevalent cases)—49 males and 153 females—were systematically ascertained through clinics within Departments of Nuclear Medicine in the Royal Marsden NHS Trust and United Medical and Dental Schools of Guy’s and St Thomas, London, between the years 1994 and 1996. These hospitals are tertiary referral centers serving the population within South England. The number of cases analyzed in this study represents ~50% of all patients under review. There was no selection of cases on the basis of family history, age, sex or histology. Blood samples from 298 healthy spouses of lung cancer cases being recruited as part of a National Cancer Research Network Trial within the UK, between the years 1999 and 2002, served as controls—91 males and 207 females. None of the controls had a personal history of malignancy. All the cases and controls were Caucasian and their ancestry was from the British Isles.

Molecular analyses

Genomic DNA was extracted from peripheral blood using conventional methods. The Q2511R polymorphism in Tg results from a single base-pair change (7589A→G), which leads to the creation of a novel TaqI restriction site. The region flanking codon 2511 was amplified using primers 5′-AATCTCGA-CAGTCGACTTCC-3′ and 5′-CGACGGGCACTCTGAGTCTC-3′ and the 240 bp amplicon was digested with a TaqI according to the manufacturer’s instructions (Promega, Madison, WI). Products were visualized on ethidium-bromide stained agarose gels.

The P52T polymorphism in TSHR results from a single base-pair change (254C→A) creating a novel Thr111 restriction site. The D276E polymorphism results from a single base pair change (2281C→G) creating a novel NlaIII restriction site. Detection of these polymorphisms was undertaken according to the methods described by Allahbadi et al. (25) and Gabriel et al. (24), respectively.

Statistical analyses

The relationship between Tg, and TSHR genotypes and risk of NMTC was assessed by means of the odds ratio (OR) with 95% confidence limits calculated by either conditional (study group 1) or unconditional logistic regression (study group 2, adjusting for sex and age). A test for trend (P trend) in increasing the risk of NMTC by having more than one putative high-risk allele of Tg was also evaluated. To test for population stratification, the distribution of genotypes in cases and controls was tested for a departure from Hardy–Weinberg equilibrium by means of the χ² test. Pooled estimates of the OR for both studies were obtained by calculating a weighted-average of the logarithm of ORs (26).

The power of each of the studies to demonstrate an association was estimated using the method published by Fleiss et al. (27), implemented in the statistical program POWER (Epicenter Software, Version 1.30; http://icarus2.hsc.usc.edu/epicenter/). The population attributable fraction was estimated by \((\chi^2 - 1)/\chi^2\), where \(\chi^2 = (1 - p)^2 + 2p(1 - \frac{OR_1}{OR}) + \frac{OR_1}{OR_2} p^2\delta\), given population allele frequency \(p\) and ORs, and \(OR_1\) and \(OR_2\) are the ORs associated with hetero- and homozygosity, respectively.

---

Table I. Demographic background of cases and controls in the two study populations

<table>
<thead>
<tr>
<th>Study group 1, Canadian</th>
<th>Study group 2, British Isles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cases</strong></td>
<td><strong>Controls</strong></td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>30/72</td>
</tr>
<tr>
<td>Age, years (±SD)</td>
<td>50.9 (23-92)</td>
</tr>
<tr>
<td>Histology</td>
<td>Papillary</td>
</tr>
<tr>
<td></td>
<td>Follicular</td>
</tr>
<tr>
<td></td>
<td>Hurtle cell</td>
</tr>
<tr>
<td></td>
<td>Anaplastic</td>
</tr>
<tr>
<td>Mixed follicular and papillary</td>
<td>14</td>
</tr>
<tr>
<td>Mixed follicular and Hurtle cell</td>
<td>2</td>
</tr>
<tr>
<td>Mixed papillary, follicular, and Hurtle cell</td>
<td>1</td>
</tr>
<tr>
<td>Unspecified</td>
<td>1</td>
</tr>
<tr>
<td><strong>Ethnicity(^a)</strong></td>
<td><strong>Cases</strong></td>
</tr>
<tr>
<td>British Isles</td>
<td>–</td>
</tr>
<tr>
<td>English Canadian and British</td>
<td>12</td>
</tr>
<tr>
<td>French Canadian and French</td>
<td>14</td>
</tr>
<tr>
<td>Jewish</td>
<td>29</td>
</tr>
<tr>
<td>Non-British and non-French European</td>
<td>37</td>
</tr>
<tr>
<td>Others</td>
<td>7</td>
</tr>
</tbody>
</table>

\(^a\)Three cases from Canada were of mixed ancestry: two half French-Canadian, half Italian; one half British, one-quarter French, one-quarter German; one control was half British, half Swiss.
Results

The frequency of the TSHR-52T allele in cases and controls in the two studies was 0.06 and 0.07, respectively; the frequency of the TSHR-727E allele was 0.05 in cases and 0.07 in controls. Because of the very small numbers of variant homozygotes observed for these polymorphisms, in estimating ORs we combined data on heterozygotes and variant homozygotes. There was no significant difference in allele frequency of the TSHR P52T and D727E polymorphisms between cases and controls in either of the two study groups (Table II). Furthermore, there was no evidence for an association from the pooled analysis. Despite their close physical proximity, there was no evidence from either study that the alleles at the two TSHR loci are in linkage disequilibrium (data not shown).

Table II also details the frequencies of Tg genotypes in cases and controls in the two study groups and the pooled analysis of the ORs from the two studies. The distribution of Tg Q2511R genotypes in cases and controls in both study groups were in Hardy–Weinberg equilibrium. The frequency of Tg-2511R allele was higher in cases than controls in both study groups (0.55 versus 0.51, and 0.55 versus 0.47 in study groups 1 and 2, respectively). These differences in allele frequency were reflected in the ORs associated with Q2511R hetero- and homozygosity, although statistical significance was only reached in study group 2. Pooling data from analysis of the two groups the ORs associated with hetero- and homozygosity were 1.6 (95% CI: 1.1–2.5) and 2.0 (95% CI: 1.2–3.3), respectively.

Discussion

Previous studies have shown that variants in Tg and TSHR (see OMIM 188450 and 603372 for a review) are associated with an increased risk of a number of benign thyroid disorders. Benign thyroid diseases such as MNG are strong risk factors for NMTC, and mutations in TSHR have been linked directly to tumor formation (20,21,28–30). Such mutations usually lead to constitutive activation of the receptor signaling activity, which in turn presumably results in uncontrolled growth of the cells. We therefore reasoned that known polymorphisms in these two genes might be associated with an increased risk for NMTC.

Ours is the first allelic association study of NMTC. We did not find any evidence of increased risk of NMTC associated with TSHR variants, P52T and D727E. The first polymorphism we examined, P52T, is located in the extra-cellular domain of the protein. The variant protein is predicted to lack a beta-turn at position 52, which would potentially alter the three-dimensional conformation of the receptor. However, examination of two individuals homozygous for the variant T-allele revealed no abnormality in thyroid function (31). Another report, which observed an association between the variant D727E and MNG, found that the E variant form of TSHR, when transiently expressed in COS-7 cells, induced a greater cAMP response to TSH stimulation than its wild-type counterpart (24). However, in vitro expression of both forms of the protein in TSA-201 cells by another group revealed no difference in TSHR activity upon TSH stimulation (32).

Our study does not, however, preclude the possibility of an association as both study groups had <40% power to demonstrate a relationship if the risk associated with carrier status was ~1.5. Furthermore, the combined dataset has only 54% power to detect an association. To evaluate these variants as possible determinants of NMTC susceptibility will require additional sample sets of sufficient size commensurate with the detection of small genotypic risks.

Levels of circulating Tg have been reported to be predictive of NMTC development suggesting the possibility of a causal relationship (33). The frequency of the R-allele of TgQ2511R was over represented in cases with NMTC in both study groups compared with controls. Only in the UK-based analysis was the allele frequency statistically different between cases and controls. Results from the Canadian study, although showing the same trend, were not statistically significant. Although we cannot entirely exclude the possibility of a type 1 error in the UK study, the power to detect a genotypic risk of 1.5 in the Canadian study is 25%. Hence there is a distinct possibility that the failure to demonstrate an association is a consequence of inadequate power. We have attempted to address this through a meta-analysis of the risk estimates generated in the two studies. Pooling data the risk of NMTC associated with

Table II. Case-control comparison of Tg and TSHR genotypes in the two study populations

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Study group 1, Canadian</th>
<th>Study group 2, British Isles</th>
<th>Pooled OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Tg Q2511R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q/Q</td>
<td>17</td>
<td>27</td>
<td>1.0</td>
</tr>
<tr>
<td>Q/R</td>
<td>57</td>
<td>46</td>
<td>2.0 (1.0–4.1)</td>
</tr>
<tr>
<td>R/R</td>
<td>28</td>
<td>29</td>
<td>1.5 (0.7–3.3)</td>
</tr>
<tr>
<td>F/P</td>
<td>93</td>
<td>87</td>
<td>1.0</td>
</tr>
<tr>
<td>P/T and T/T</td>
<td>15</td>
<td>15</td>
<td>0.6 (0.2–1.4)</td>
</tr>
<tr>
<td>TSHR D727E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D/D</td>
<td>88</td>
<td>88</td>
<td>1.0</td>
</tr>
<tr>
<td>D/E and E/E</td>
<td>14</td>
<td>14</td>
<td>1.0 (0.4–2.4)</td>
</tr>
</tbody>
</table>

*Adjusted for age and sex.

Also shown are the pooled estimates of ORs associated with genotypes.
hetero- and homozygosity for the R-allele were 1.6 and 2.0, respectively. The mechanism by which this polymorphism may confer an increased risk of NMTC remains to be established. While the TgQ2511R change might mediate risk through linkage disequilibrium with a functional variant, the glutamine to arginine amino acid substitution is non-conservative and resides within the acetylcholinesterase homologous domain of TG. The putative function of this domain is not yet clear, but as acetylcholinesterase interacts with cell membranes in the nervous system, a similar role for the homologous domain at the C-terminal end of the TG molecule has been proposed for apical membrane interaction (34). There is direct evidence for functionality of this domain in man with missense mutations being associated with fetal goitre (35). Furthermore, mutations in this domain in both mouse and rat model systems has been shown to cause storage of apparently misfolded TG molecules in the endoplasmic reticulum leading to congenital goiter and hypothyroidism (36,37).

Taken together the data suggest that germline variation in Tg might either exert a direct effect on NMTC risk or mediate risk through predisposition to goiter. Although the risk of NMTC associated with the R-allele is modest, the high prevalence of the allele in the general population means that it is likely to make a significant impact on NMTC cancer incidence. On the basis that the allele frequency is ~50% in the Caucasian population the risk in hetero- and homozygotes translates to this variant underlying around one-third of all cases. A number of other variants in Tg have been documented. Given our finding it is possible that other variants in the gene may confer an increased risk either independently or by defining a high-risk haplotype. Further studies will be required to confirm this finding and further investigate the precise role of the Tg gene in NMTC pathogenesis.

Acknowledgements

We thank Elsa Lanke for technical assistance. This work was funded by the Fonds de la recherché en Sante du Quebec (FRSQ), the Canadian Genetic Diseases Network, the Thyroid Foundation of Canada (W.D.F.) and Cancer Research UK. W.D.F. is a Chercheur Clinicien of the FRSQ. A.M. was in receipt of a clinical research fellowship from the Allan J. Learner Fund.

References


Received July 2, 2003; revised October 15, 2003; accepted November 11, 2003