Published in final edited form as:

Nat Genet. ; 44(3): 319–322. doi:10.1038/ng.1046.

Discovery of common variants associated with low TSH levels and thyroid cancer risk

Julius Gudmundsson^{1,16}, Patrick Sulem^{1,16}, Daniel F Gudbjartsson¹, Jon G Jonasson^{2,3,4}, Gisli Masson¹, Huiling He⁵, Aslaug Jonasdottir¹, Asgeir Sigurdsson¹, Simon N Stacey¹, Hrefna Johannsdottir¹, Hafdis Th Helgadóttir¹, Wei Li⁵, Rebecca Nagy⁵, Matthew D Ringel⁶, Richard T Kloos⁶, Marieke C H de Visser⁷, Theo S Plantinga⁸, Martin den Heijer^{7,15}, Esperanza Aguillo⁹, Angeles Panadero¹⁰, Enrique Prats¹¹, Almudena Garcia-Castaño¹², Ana De Juan¹², Fernando Rivera¹², G Bragi Walters¹, Hjördis Bjarnason¹, Laufey Tryggvadóttir^{3,4}, Gudmundur I Eyjolfsson¹³, Unnur S Björnsdóttir³, Hilma Holm¹, Isleifur Olafsson², Kristleifur Kristjánsson¹, Hoskuldur Kristvinsson², Olafur T Magnusson¹, Gudmar Thorleifsson¹, Jeffrey R Gulcher¹, Augustine Kong¹, Lambertus ALM Kiemeny⁷, Thorvaldur Jonsson^{2,3}, Hannes Hjartarson², Jose I Mayordomo¹⁴, Romana T Netea-Maier¹⁵, Albert de la Chapelle⁵, Jon Hrafnkelsson², Unnur Thorsteinsdóttir^{1,3}, Thorunn Rafnar¹, and Kari Stefansson^{1,3}

¹deCODE Genetics, Reykjavik, Iceland ²Lanspítali University Hospital, Reykjavik, Iceland ³Faculty of Medicine, University of Iceland, Reykjavik, Iceland ⁴The Icelandic Cancer Registry, Reykjavik, Iceland ⁵Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio, USA ⁶Division of Endocrinology, Diabetes, and Metabolism, The Ohio State University, Columbus, Ohio, USA ⁷Department of Epidemiology, Biostatistics and Health Technology Assessment (HTA), Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands ⁸Department of Internal Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands ⁹Division of Endocrinology, University Hospital, Zaragoza, Spain ¹⁰Division of Medical Oncology, Ciudad de Coria Hospital, Coria, Spain ¹¹Division of Nuclear Medicine, University Hospital, Zaragoza, Spain ¹²Division of Medical Oncology, Marques de Valdecilla University Hospital, Santander, Spain ¹³The Laboratory in Mjódd, Reykjavik, Iceland ¹⁴Division of Medical Oncology, University of Zaragoza, Zaragoza, Spain ¹⁵Department of Endocrinology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Abstract

Correspondence should be addressed to: K.S. (kstefans@decode.is) or J.G. (julius.gudmundsson@decode.is).

¹⁶These authors contributed equally to this work.

Accession codes. Gene Expression Omnibus (GEO) database: blood and adipose tissue samples, GSE7965; human 3.0 A1, GPL3991.

Note: Supplementary information is available on the Nature Genetics website.

AUTHOR CONTRIBUTIONS

The study was designed and results were interpreted by J.G., P.S., D.F.G., A.K. and K.S. Statistical analysis was carried out by P.S., D.F.G., G.M., G.T., J.G. and A.K. Subject recruitment and biological material collection and handling was organized and carried out by J.G., J.G.J., S.N.S., H.He, W.L., R.N., M.D.R., R.T.K., M.C.H.de.V., T.S.P., M.d.H., E.A., A.P., E.P., A.G.-C., A.D.J., F.R., G.B.W., H.B., L.T., I.O., G.I.E., U.S.B., H.Holm, K.K., H.K., J.R.G., L.A.L.M.K., R.T.N.-M., T.J., H. Hjartarson, J.I.M., A.de.la.C., J.H., U.T. and T.R. Genotyping was supervised and carried out by J.G., A.J., A.S., H.He, H.J., H.Th.H., O.Th.M., W.L. and U.T. J.G., P.S., D.F.G. and K.S. drafted the manuscript. All authors contributed to the final version of the paper. Principal collaborators for the replication case-control samples were J.I.M. (Spain), A.d.l.C. (US) and R.T.N.-M. and L.A.L.M.K. (The Netherlands).

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

To search for sequence variants conferring risk of nonmedullary thyroid cancer, we focused our analysis on 22 SNPs with a $P < 5 \times 10^{-8}$ in a genome-wide association study on levels of thyroid stimulating hormone (TSH) in 27,758 Icelanders. Of those, rs965513 has previously been shown to associate with thyroid cancer. The remaining 21 SNPs were genotyped in 561 Icelandic individuals with thyroid cancer (cases) and up to 40,013 controls. Variants suggestively associated with thyroid cancer ($P < 0.05$) were genotyped in an additional 595 non-Icelandic cases and 2,604 controls. After combining the results, three variants were shown to associate with thyroid cancer: rs966423 on 2q35 (OR = 1.34; $P_{\text{combined}} = 1.3 \times 10^{-9}$), rs2439302 on 8p12 (OR = 1.36; $P_{\text{combined}} = 2.0 \times 10^{-9}$) and rs116909374 on 14q13.3 (OR = 2.09; $P_{\text{combined}} = 4.6 \times 10^{-11}$), a region previously reported to contain an uncorrelated variant conferring risk of thyroid cancer. A strong association ($P = 9.1 \times 10^{-91}$) was observed between rs2439302 on 8p12 and expression of *NRG1*, which encodes the signaling protein neuregulin 1, in blood.

The contribution of genetics to the risk of thyroid cancer is greater than to any other cancer, and the effect extends beyond the nuclear family^{1–4}. Thyroid cancer is classified into four main histology groups: papillary (PTC), follicular (FTC), medullary (MTC), and undifferentiated or anaplastic thyroid carcinomas. The great majority of malignant thyroid tumors are nonmedullary, either PTC (80–85%) or FTC (10–15%)^{5,6}.

Among sequence variants that have been implicated in the etiology of PTC are variants at 1p12, 8q24, and the pre-miR146a at 5q33 (refs. 7–10). Furthermore, in results from a genome-wide association study (GWAS) on thyroid cancer, two common variants, located on 9q22.33 and 14q13.3, were shown to associate with both PTC and FTC¹¹. The association with variants on 9q22.33 has been replicated both in sporadic¹² and radiation-induced¹³ thyroid cancers. The two variants have also been associated with low serum concentrations of thyroid-stimulating hormone (TSH), and the 9q22.33 variant associates with low concentration of free thyroxine (FT₄) and high concentration of free triiodothyronine (FT₃)¹¹. Here we focused our search for additional thyroid cancer risk variants on SNPs with $P < 5 \times 10^{-8}$ in a GWAS on circulating TSH levels.

The GWAS on TSH levels included 27,758 individuals not known to have thyroid cancer, on the basis of information from the nationwide Icelandic Cancer Registry (Fig. 1, Supplementary Table 1 and Supplementary Fig. 1). Genotyping was done using the Illumina SNP chip platform. We also made use of results from genome-wide sequencing of 457 Icelanders to an average depth of over 10× (see Fig. 1 and the Supplementary Note), resulting in the identification of some 16 million SNPs. Using imputation assisted by long-range haplotype phasing^{14,15}, we inferred the genotypes of these SNPs in 41,675 Icelanders who had been genotyped using the Illumina SNP chip platform, including the 27,758 with TSH measurements.

On the basis of an association analysis of this data set, we found 22 variants that associate with serum levels of TSH at a significance threshold of $P < 5 \times 10^{-8}$ (Supplementary Table 2). For one of these regions, 9q22.33, a variant (rs965513) associating with both thyroid cancer and TSH levels has been reported¹¹ and will not be discussed any further here. Of the remaining 21 TSH-associated loci, three associate with levels of TSH—1p36 (ref. 16), 5q14.1 (refs. 17,18) and 6q27 (ref. 17)—but none has been reported to associate with thyroid cancer.

By inspecting our GWAS data set on Icelandic nonmedullary thyroid cancer, generated with both chip and imputed genotypes (see Online Methods for a detailed description), 5 of the 21 TSH-associated SNPs were found to associate with thyroid cancer with nominal significance ($P < 0.05$, Supplementary Table 2). We also searched for stronger or additional thyroid cancer association signals by examining the thyroid cancer GWAS results for SNPs

located within a 1.5-Mb region centered on each of the 21 original TSH SNPs. On the basis of this analysis, we found a SNP (rs966423) on 2q35 with a more significant association with thyroid cancer than the initial TSH-associated SNP (rs737308) at this locus ($P = 0.0010$ and 0.31 for rs966423 and rs737308, respectively). The pairwise correlation between the two SNPs on 2q35 was low ($r^2 = 0.003$, $D' = 0.08$, according to data from 2,349 Icelanders), and therefore, they are likely to represent different association signals. rs966423 also associated with serum levels of TSH, but the effect was smaller and the significance less than for rs737308 (Supplementary Table 2). We found no additional variants associated with thyroid cancer at the remaining TSH loci.

We genotyped the 21 TSH-associated SNPs and the one suggestive thyroid cancer SNP (rs966423) on 2q35 in 561 individuals with thyroid cancers and 3,190 controls from the general Icelandic population, using the Centaurus¹⁹ single-track genotyping assay. Four of the TSH-associated SNPs are present on the Illumina chips, and for those, we made use of chip-derived genotypes from 39,864 individuals (see Online Methods for a detailed description). Analysis of directly genotyped individuals (using either the single-track assay or the Illumina chip platform) resulted in a nominally significant ($P < 0.05$) association signal for thyroid cancer and five SNPs located on 1p31.3, 1p36.13, 2q35, 8p12 and 14q13.3 (Supplementary Table 2). At 2q35, only the variant rs966423 was significantly associated with thyroid cancer, whereas rs737308, the SNP on 2q35 with a stronger association with TSH levels, was not ($P = 3.8 \times 10^{-4}$ and 0.26 for rs966423 and rs737308, respectively).

Four of the five variants were genotyped and tested for association with thyroid cancer in three case-control groups of European descent, with populations from the USA (Ohio), The Netherlands and Spain. The fifth variant, rs10799824 on 1p36.13, was genotyped only in the Dutch and Spanish samples, but the association was not replicated, and this SNP was not studied further. The results for the variant on 1p31.3 (rs334725) did replicate in two out of the three case-control groups, resulting in a combined allelic odds ratio (OR) of 1.31 ($P = 6.6 \times 10^{-3}$; Supplementary Table 3). Whether this variant truly confers risk of thyroid cancer remains to be shown. Combining the results from Iceland and the follow-up groups gave OR estimates between 1.34 and 2.09 for the remaining three variants located on 2q35, 8p12 and 14q13.3 ($P < 3 \times 10^{-9}$; Table 1).

Of the three genome-wide significant thyroid cancer variants reported here, the strongest association was observed for allele T of rs116909374 (rs116909374[T]) on 14q13.3. This SNP and the SNP rs944289, which has been reported to associate with thyroid cancer¹¹, are located within two distinct but neighboring LD regions (Supplementary Fig. 2). The correlation between them is very low ($r^2 = 0.005$, $D' = 0.35$, according to data from 3,693 Icelanders), and the association with thyroid cancer for each SNP remains significant after adjusting for the other (Supplementary Table 4). The association effect for TSH levels is substantially stronger for the current SNP than for the previously reported one (effect = 0.141 s.d. and $P = 1.1 \times 10^{-16}$ for rs116909374[T] compared to an effect = 0.022 s.d. and $P = 0.001$ for rs944289[T]). This suggests that the 14q13.3 locus contains more than one variant causing a predisposition to thyroid cancer or, alternatively, that a unique variant capturing the effect of rs116909374 and rs944289 remains to be discovered. On 14q13.3, the gene closest to rs116909374 is *MBIP* (a regulatory protein), but another nearby candidate gene, which must be considered because of its prominent role in the development of the thyroid²⁰, is the thyroid transcription factor *NKX2-1*. The variant rs966423, at 2q35, is located in the *DIRC3* gene and has not previously been associated with thyroid cancer or serum levels of TSH.

The SNP rs2439302, on 8p12, is located within the first intron of the gene *NRG1*. The *NRG1* gene encodes neuregulin 1, a signaling protein that mediates cell-cell interactions and

plays an important role in the development of the nervous system, heart, breast and other organs. Germline sequence variation at the *NRG1* locus has been associated with schizophrenia²¹ and with Hirschsprung's disease in individuals of Asian descent²². The thyroid cancer variant reported here (rs2439302[G]) and one of the two originally reported variants (rs7835688[C]) for the Hirschsprung's disease are positively correlated on the basis of HapMap data from both Asian (CHB+JPT; $r^2 = 0.95$) and European-descended (CEU; $r^2 = 0.94$) individuals. Hence, it is likely that the same sequence variant in *NRG1* affects circulating TSH levels and risks of Hirschsprung's disease and thyroid cancer.

To determine whether any of the three thyroid cancer variants associate with expression levels of the genes in which they are either located or close by, we examined our previously described microarray expression data set²³ from whole blood of 966 population-based Icelandic controls. No significant association was detected between the variants on 2q35 and 14q13.3, and expression levels of genes located within a 1-Mb region, centered on the thyroid cancer risk variants. However, a highly significant association was observed between rs2439302, on 8p12, and expression measured by several different probes in the *NRG1* gene, the only Reference Sequence (RefSeq) database gene within a 1-Mb region centered on rs2439302. Carriers of the allele conferring risk of thyroid cancer (rs2439302[G]) have a lower relative expression of *NRG1*. For the probe with the most significant association ($P = 9.1 \times 10^{-91}$), the relative expression of *NRG1* was decreased by 40% per each allele G of rs2439302. Hence, the relative expression of *NRG1* is 64% lower in homozygous carriers of the thyroid cancer risk allele than in homozygous noncarriers (Fig. 2). The strong association results seen for rs2439302 and thyroid cancer on one hand, and the expression levels of *NRG1* on the other hand, strongly suggest that *NRG1* plays a role in the etiology of nonmedullary thyroid cancer.

Using only the Icelandic study group, no significant effect was detected for age at diagnosis for any of the variants reported here. In terms of effect of the new variants on the histological subclasses of thyroid cancer, the number of FTC samples were, in general, too few to draw meaningful conclusions.

For the SNPs rs966423, rs2439302 and rs116909374, the sibling recurrence risk ratio was estimated to be 1.021, 1.023 and 1.033, respectively, on the basis of combined effect and combined frequencies from all four study groups. In order to summarize the overall effect of these three variants and the two previously reported variants on 9q22.33 and 14q12.3, we combined the effect of all five variants known today to confer risk of thyroid cancer. We used the combined risk estimates for the four study population used in the present study, assuming a multiplicative model for the five risk variants. On the basis of this analysis, the estimated risk of the disease was over 2.3 times greater for the top 5% of the risk distribution, than for the general population. Similarly, the risk of the disease was over three times greater for the top 1% of the risk distribution. We note that the estimates provided here are based solely on populations used in this study and will have to be updated as new variants are discovered.

For the 22 variants associating with TSH levels at a significance threshold of $P < 5 \times 10^{-8}$, the fraction of variance of TSH explained was 4.3%. We also checked their association with risk of goiter and with levels of FT₄ and FT₃. Of the 22 TSH loci, 9 associated with risk of goiter with nominal significance ($P < 0.05$). Two of these loci, 1p36.13 and 15q21.1, were recently reported in a GWAS on goiter risk²⁴. Two of the TSH variants had a nominally significant association with levels of FT₃ (rs965513 on 9q22.33 and rs61938844 on 12q23.1; $P < 0.05$), and 12 of the 22 TSH variants had a nominally significant association with levels of FT₄.

Taken together, these results demonstrate the unresolved complexity of DNA polymorphism and their effects on biological functions such as levels of thyroid-related hormones and risk of both benign and malignant thyroid diseases. The variants associating with serum TSH concentrations can be divided into two main groups; those that do and those that do not confer risk of thyroid cancer (Supplementary Fig. 3). Notably, for all the variants that confer risk of thyroid cancer, the at-risk allele for thyroid cancer associated with lower serum levels of TSH. This was true for the three variants reported here as well as for the two thyroid cancer-TSH variants previously published¹¹. Furthermore, the variants on 2q35 and 14q13.3 reported above also associated with increased levels of FT₄. Hence, it appears that for carriers of these five variants the primary disorder in nonmedullary thyroid cancer is an endocrine one, characterized by a lower concentration of TSH. The consequence of the low concentration of TSH may be less differentiation of the thyroid epithelium, leading to a predisposition to malignant transformation. There is, however, more to the story of thyroid cancer because not all sequence variants associating with low TSH levels associate with the risk of the disease.

URLs

The Icelandic Cancer Registry, <http://www.krabbameinsskra.is/indexen.jsp?icd=C73>.

ONLINE METHODS

Icelandic study population

All participants in this study are of European ancestry. Individuals diagnosed with thyroid cancer were identified based on a nationwide list from the Icelandic Cancer Registry (ICR) that contains all Icelanders with thyroid cancer diagnosed from January 1, 1955, to December 31, 2009. Of these, 1,018 were nonmedullary thyroid cancers. The histology of all thyroid carcinomas used in the present study has been reviewed and confirmed by one of the authors of this article (J.G.J.). Included in the present study are DNA samples from 572 individuals with nonmedullary thyroid cancer, diagnosed from December 1974 to December 2009 and who were recruited from November 2000 until April 2010. The median time from diagnosis to blood sampling is 10 years (range 0–46 years). The mean age at diagnosis of those recruited was 44 years (median 43 years), and the range was 13–87 years, whereas the mean age at diagnosis is 56 years for all individuals with thyroid cancer in the ICR. The thyroid cancer GWAS data set used in the current study comprises results from 222 individuals with thyroid cancer and 24,198 controls genotyped using Illumina Human Hap300, HapCNV370, Hap610, 1M, or Omni-1 Quad BeadChips (Illumina), as well as results from 627 individuals with thyroid cancer and 71,613 unaffected individuals with genotypes inferred using an imputation method making use of the Icelandic genealogy to propagate genotypic information into individuals for whom we have neither SNP chip nor sequence data, a process we refer to as genealogy-based imputation, the combined method of imputing sequence-derived data into Illumina chip-typed individuals and using genealogy-based imputation to infer the DNA sequence of ungenotyped individuals as two-way imputation (see Supplementary Note). For confirming thyroid cancer GWAS results for SNPs under investigation in the present study, we used the Centaurs genotyping platform (see Supplementary Note) to attempt genotyping all 572 samples available from affected individuals and a minimum of 1,500 unaffected individuals. Of these, 561 samples from affected individuals and a minimum of 1,472 unaffected individuals (~98%) were successfully genotyped in our study. Of the 561 affected individuals genotyped using the Centaurus platform, 222 had previously been genotyped using the Illumina chips (see Supplementary Note). The data overlap was used to confirm data consistency. The remaining 339 affected individuals genotyped using the Centaurus platform are a subset of the 627 affected individuals contributing imputed genotypes to the initial thyroid cancer

GWAS data set (Supplementary Table 1b). Individuals with goiter were defined as those with SNOMED codes 71600, 71602, 71620 or 71640 in the pathology registries of the Landspítali University Hospital in Reykjavik and The Regional Hospital, Akureyri, Iceland. The period from which data was collected was from 1982 to 2010. Our total list of individuals with goiter consists of 691 individuals, of whom 217 have been genotyped using the Illumina SNP chip platform and 329 have their genotypes imputed using two-way imputation based on first- or second-degree SNP chip-genotyped relatives. The 40,013 unaffected individuals (17,326 males (43.3%) and 22,687 females (56.7%)) used in this study consisted of individuals belonging to different genetic research projects at deCODE. The unaffected individuals had a mean age of 61 years (s.d. = 20.6 years). The unaffected individuals were absent from the nationwide list of individuals with thyroid cancer, according to the ICR. The DNA for both the Icelandic groups was isolated from whole blood using standard methods. The study was approved by the Data Protection Commission of Iceland and the National Bioethics Committee of Iceland. Written informed consent was obtained from all persons. Personal identifiers associated with medical information and blood samples were encrypted with a third-party encryption system as previously described²⁵.

The Netherlands (Dutch) study population

Participants consisted of 151 individuals with nonmedullary thyroid cancer (75% females) and 832 cancer-free individuals (54% females). Individuals were recruited from the Department of Endocrinology, Radboud University Nijmegen Medical Centre (RUNMC), Nijmegen, The Netherlands from November 2009 to June 2010. All affected individuals were of self-reported European descent. Demographic, clinical, tumor treatment and follow-up related characteristics were obtained from the medical records of affected individuals. The average age at diagnosis was 39 years (s.d. = 12.8). The DNA for both the affected and unaffected individuals was isolated from whole blood, with standard methods. The unaffected individuals were recruited within a project entitled Nijmegen Biomedical Study (NBS). The details of this study have been reported previously²⁶. Unaffected individuals from the NBS were invited to participate in a study on gene-environment interactions in multifactorial diseases such as cancer. They were all of self-reported European descent and fully informed about the goals and the procedures of the study. The study was approved by the Ethical Committee and the Institutional Review Board of the RUNMC, Nijmegen, The Netherlands and all study subjects gave written informed consent.

USA study population

All individuals were of self-reported European descent and provided written informed consent. These individuals with thyroid cancer ($n = 365$; median age 40 years, range 13–80 years; 76% are females) were recruited from Ohio, USA, and were histologically confirmed as having papillary thyroid carcinoma (PTC) (including traditional PTC and follicular variant PTC). The control group ($n = 383$; median age 49 years, range 18–87 years; 65% are females) comprised individuals without clinically diagnosed thyroid cancer from the central Ohio area. Genomic DNA was extracted from blood. The study was approved by the Institutional Review Board of Ohio State University.

Spain study population

The Spanish study population consisted of 90 individuals with nonmedullary thyroid cancer. They were recruited from the Oncology Department of Zaragoza Hospital in Zaragoza, Spain, from October 2006 to June 2007. All were of self-reported European descent. Clinical information including age at onset, grade and stage was obtained from medical records. The average age at diagnosis for the affected individuals was 48 years (median 49 years), and the range was 22–79 years. The 1,399 Spanish unaffected individuals 798 (57%)

males and 601 (43%) females had a mean age of 51 (median age 50 and range 12–87 years) were approached at the University Hospital in Zaragoza, Spain, and were not known to have thyroid cancer. The DNA for both the affected and unaffected individuals was isolated from whole blood using standard methods. Study protocols were approved by the Institutional Review Board of Zaragoza University Hospital. All individuals in the study gave written informed consent.

TSH, FT₄ and FT₃ measurements

TSH, FT₄ and FT₃ levels were measured in blood samples of Icelanders seeking medical care between the years 1997 and 2010 at the Landspítali University Hospital or at the Clinical Laboratory in Mjódd, Reykjavik, Iceland. Measurements outside the specified range were discarded (Supplementary Table 1a). The measurements were normalized to a standard normal distribution using quantile-quantile normalization and then adjusted for center, gender, year of birth and age at measurement. For individuals for which more than one measurement was available, we used the average of the normalized value. The normalized and adjusted measurements were regressed on allele counts using classical linear regression.

Expression analysis

The expression analysis has been described previously²³. Briefly, the subjects were randomly selected and recruited as dense three-generation families as a part of a study on inheritance of gene expression in humans. Peripheral blood was collected and DNA and RNA extracted. The RNA samples were hybridized to a custom made microarray (Agilent Technologies). The correlation between the mean logarithm ($\log_{10}$) expression ratio (MLR) of *NRG1*, measured using RNA from whole blood of 966 individuals, and genotypes of relevant SNP(s), was tested by regressing the MLRs on the number of copies of the at-risk alleles. The effect of age, gender and differential cell counts was taken into account by including corresponding terms as an explanatory variable in the regression analysis. All *P* values were adjusted for relatedness of the individuals by dividing the chi-square statistics by an adjustment factor of 1.064, determined through simulations. The probe used to test the expression of *NRG1* is located on 8p12 at the position chr. 8:32741415–32741475 in NCBI build 36. The total expression data set is available from the GEO database (see accession code listings above).

Genotyping

The Icelandic chip-typed samples were assayed using the Illumina BeadChips at deCODE genetics. Single-track assay SNP genotyping of the SNPs reported in Table 1 of the main text for the three case-control groups from Iceland, The Netherlands and Spain was carried out by deCODE Genetics, applying the Centaurus²⁷ (Nanogen) platform or the Illumina SNP chips. Genotyping of samples from the Ohio study populations was done using the SNaPshot (PE Applied Biosystems) genotyping platform at the Ohio State University, as previously described²⁸.

Association testing

Logistic regression was used to test for association between SNPs and disease, treating disease status as the response and expected genotype counts from imputation or allele counts from direct genotyping as covariates. Testing was performed using the likelihood ratio statistic^{29–31}. For quantitative trait association testing, the normalized and adjusted measurements of TSH, FT₃ and FT₄ were regressed on allele counts using classical linear regression. Information on genealogy-based imputation, the Centaurus genotyping platform and likelihood ratio statistical testing (including a detailed description of the genotyping, sequencing and statistical methods) is available in the Supplementary Note.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the affected individuals whose contribution made this work possible. This project was funded in part by US National Institutes of Health contract numbers CA16058 and CA124570.

References

1. Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst.* 1994; 86:1600–1608. [PubMed: 7932824]
2. Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. *Int J Cancer.* 2002; 99:260–266. [PubMed: 11979442]
3. Amundadottir LT, et al. Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS Med.* 2004; 1:e65. [PubMed: 15630470]
4. Hrafnkelsson J, Tulinius H, Jonasson JG, Sigvaldason H. Familial non-medullary thyroid cancer in Iceland. *J Med Genet.* 2001; 38:189–191. [PubMed: 11303513]
5. Kondo T, Ezzat S, Asa SL. Pathogenetic mechanisms in thyroid follicular-cell neoplasia. *Nat Rev Cancer.* 2006; 6:292–306. [PubMed: 16557281]
6. DeLellis RA. Pathology and genetics of thyroid carcinoma. *J Surg Oncol.* 2006; 94:662–669. [PubMed: 17131411]
7. Baida A, et al. Strong association of chromosome 1p12 loci with thyroid cancer susceptibility. *Cancer Epidemiol Biomarkers Prev.* 2008; 17:1499–1504. [PubMed: 18559567]
8. Jazdzewski K, et al. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proc Natl Acad Sci USA.* 2008; 105:7269–7274. [PubMed: 18474871]
9. Jazdzewski K, et al. Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. *Proc Natl Acad Sci USA.* 2009; 106:1502–1505. [PubMed: 19164563]
10. He H, et al. A susceptibility locus for papillary thyroid carcinoma on chromosome 8q24. *Cancer Res.* 2009; 69:625–631. [PubMed: 19147577]
11. Gudmundsson J, et al. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nat Genet.* 2009; 41:460–464. [PubMed: 19198613]
12. Landa I, et al. The variant rs1867277 in *FOXE1* gene confers thyroid cancer susceptibility through the recruitment of USF1/USF2 transcription factors. *PLoS Genet.* 2009; 5:e1000637. [PubMed: 19730683]
13. Takahashi M, et al. The *FOXE1* locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl. *Hum Mol Genet.* 2010; 19:2516–2523. [PubMed: 20350937]
14. Kong A, et al. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat Genet.* 2008; 40:1068–1075. [PubMed: 19165921]
15. Kong A, et al. Fine-scale recombination rate differences between sexes, populations and individuals. *Nature.* 2010; 467:1099–1103. [PubMed: 20981099]
16. Panicker V, et al. A locus on chromosome 1p36 is associated with thyrotropin and thyroid function as identified by genome-wide association study. *Am J Hum Genet.* 2010; 87:430–435. [PubMed: 20826269]
17. Arnaud-Lopez L, et al. Phosphodiesterase 8B gene variants are associated with serum TSH levels and thyroid function. *Am J Hum Genet.* 2008; 82:1270–1280. [PubMed: 18514160]
18. Medici M, et al. A large-scale association analysis of 68 thyroid hormone pathway genes with serum TSH and FT₄ levels. *Eur J Endocrinol.* 2011; 164:781–788. [PubMed: 21367965]
19. Kutayav IV, et al. A novel endonuclease IV post-PCR genotyping system. *Nucleic Acids Res.* 2006; 34:e128. [PubMed: 17012270]

20. Parlato R, et al. An integrated regulatory network controlling survival and migration in thyroid organogenesis. *Dev Biol.* 2004; 276:464–475. [PubMed: 15581879]
21. Stefansson H, et al. Neuregulin 1 and susceptibility to schizophrenia. *Am J Hum Genet.* 2002; 71:877–892. [PubMed: 12145742]
22. Garcia-Barcelo MM, et al. Genome-wide association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. *Proc Natl Acad Sci USA.* 2009; 106:2694–2699. [PubMed: 19196962]
23. Emilsson V, et al. Genetics of gene expression and its effect on disease. *Nature.* 2008; 452:423–428. [PubMed: 18344981]
24. Teumer A, et al. Genome-wide association study identifies four genetic loci associated with thyroid volume and goiter risk. *Am J Hum Genet.* 2011; 88:664–673. [PubMed: 21565293]
25. Gulcher JR, Kristjansson K, Gudbjartsson H, Stefansson K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur J Hum Genet.* 2000; 8:739–742. [PubMed: 11039572]
26. Wetzels JF, Kiemeneij LA, Swinkels DW, Willems HL, den Heijer M. Age- and gender-specific reference values of estimated GFR in Caucasians: the Nijmegen Biomedical Study. *Kidney Int.* 2007; 72:632–637. [PubMed: 17568781]
27. Kutayvin IV, et al. A novel endonuclease IV post-PCR genotyping system. *Nucleic Acids Res.* 2006; 34:e128. [PubMed: 17012270]
28. He H, et al. Allelic variation in gene expression in thyroid tissue. *Thyroid.* 2005; 15:660–667. [PubMed: 16053381]
29. Sulem P, et al. Identification of low-frequency variants associated with gout and serum uric acid levels. *Nat Genet.* 2011; 43:1127–1130. [PubMed: 21983786]
30. Rafnar T, et al. Mutations in *BRIP1* confer high risk of ovarian cancer. *Nat Genet.* 2011; 43:1104–1107. [PubMed: 21964575]
31. Stacey SN, et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet.* 2011; 43:1098–1103. [PubMed: 21946351]

Figure 1. Study design and results. The number of individuals with thyroid cancer (affected) or unaffected individuals imputed or successfully genotyped is shown for each study segment.

Figure 2.

Correlation between relative expression of *NRG1* in blood (*y* axis) and genotypes (*x* axis) of rs2439302 on 8p12. The expression of *NRG1* is shown as 10^{MLR} , where MLR is the mean log expression ratio per individual. In the plot, the bottoms and tops of the boxes indicate the 25th and 75th percentiles, respectively, and the whiskers indicate the minimum and maximum values. The medians are shown as horizontal lines inside the boxes.

Association results for three variants on 2q23, 8p12 and 14q13.3 and thyroid cancer in Iceland, the Netherlands, Spain and the United states^a

Table 1

Study population (<i>n</i> cases/ <i>n</i> controls)	OR ^b	95% CI ^b	<i>P</i> value ^{b,c}	Affected (freq.)	Unaffected (freq.)
rs966423[C] on 2q35^d					
Iceland (546/38,854) ^d	1.26	(1.11, 1.43)	3.8 × 10 ⁻⁴	0.499	0.442
The Netherlands (149/814)	1.80	(1.40, 2.31)	4.2 × 10 ⁻⁶	0.554	0.408
United States (365/383)	1.36	(1.11, 1.67)	3.5 × 10 ⁻³	0.471	0.396
Spain (90/1,397)	1.20	(0.89, 1.62)	0.24	0.450	0.406
All combined (1,150/41,448) ^e	1.34	(1.22, 1.47)	1.3 × 10 ⁻⁹	–	–
<i>R</i> _{het}		0.079			
<i>P</i>		0.55			
rs2439302[G] on 8p12^f					
Iceland (532/3,094) ^d	1.41	(1.23, 1.62)	1.3 × 10 ⁻⁶	0.535	0.449
The Netherlands (149/806)	1.24	(0.97, 1.60)	0.088	0.520	0.466
United States (365/383)	1.33	(1.08, 1.63)	6.1 × 10 ⁻³	0.547	0.475
Spain (88/1,342)	1.34	(0.97, 1.85)	0.073	0.420	0.351
All combined (1,134/5,625) ^e	1.36	(1.23, 1.50)	2.0 × 10 ⁻⁹	–	–
<i>R</i> _{het}		0.85			
<i>P</i>		0.0			
rs116909374[T] on 14q13.3					
Iceland (542/3,190) ^d	2.03	(1.54, 2.67)	5.4 × 10 ⁻⁷	0.085	0.044
The Netherlands (151/824)	1.95	(1.09, 3.48)	0.024	0.056	0.030
United States (356/374)	1.98	(1.12, 3.49)	0.018	0.049	0.025
Spain (89/952)	3.37	(1.53, 7.44)	2.6 × 10 ⁻³	0.056	0.017
All combined (1,138/5,340) ^e	2.09	(1.68, 2.60)	4.6 × 10 ⁻¹¹	–	–
<i>R</i> _{het}		0.67			
<i>P</i>		0.0			

^a SNPs directly genotyped with either single-track assay or Illumina chip platform in cases and controls to determine allelic frequencies of risk variants in affected (case) and control individuals.

^b Values based on the multiplicative model.

^cAll P values are two sided.

^dFor rs966423, which is present on the Illumina chips and was used to genotype the Icelandic GWAS population, results are included for chip-genotyped individuals in addition to single-track assay genotyped individuals; for the other two SNPs, the results for all study groups are based on single-track assay genotyping.

^eFor the combined study populations, the OR and the P value were estimated using the Mantel-Haenszel model.

^fNote that rs2439302 is a G/C-SNP, and the coding of the alleles here is as on the plus (+) strand of the human reference sequence in build 36. OR, allelic odds ratio; CI, confidence interval; P_{het} , P value for the heterogeneity between the different study groups; $\bar{P}^2$, heterogeneity statistic representing the fraction of variability due to heterogeneity between study groups.