Identification of NID1 as a novel candidate susceptibility gene for familial non-medullary thyroid carcinoma using whole-exome sequencing

Luis Eduardo Barbalho de Mello1,2,*, Thaise Nayane Ribeiro Carneiro1,*, Aline Neves Araujo1, Camila Xavier Alves2, Pedro Alexandre Favoretto Galante3, Vanessa Candiotti Buzatto3, Maria das Graças de Almeida2,4, Karina Marques Vermeulen-Serpa3, Sancha Helena de Lima Vale4,5, Fernando José de Pinto Paiva2, José Brandão-Neto2 and Janete Maria Cerutti1

1Genetic Bases of Thyroid Tumors Laboratory, Division of Genetics, Department of Morphology and Genetics, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, São Paulo, Brazil
2Postgraduate Program in Health Sciences, Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte, Brazil
3Centro de Oncologia Molecular, Hospital Sírio-Libanés, São Paulo, São Paulo, Brazil
4Department of Clinical and Toxicological Analyses, Natal, Rio Grande do Norte, Brazil
5Department of Nutrition, Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte, Brazil

Correspondence should be addressed to J M Cerutti: j.cerutti@unifesp.br

(LEB de Mello and TN R Carneiro contributed equally to this work)

Abstract

The genetics underlying non-syndromic familial non-medullary thyroid carcinoma (FNMTCT) is still poorly understood. To identify susceptibility genes for FNMTCT, we performed whole-exome sequencing (WES) in a Brazilian family affected by papillary thyroid carcinoma (PTC) in three consecutive generations. WES was performed in four affected and two unaffected family members. Manual inspection in over 100 previously reported susceptibility genes for FNMTCT showed that no variants in known genes co-segregated with disease phenotype in this family. Novel candidate genes were investigated using PhenoDB and filtered using Genome Aggregation (gnomAD) and Online Archive of Brazilian Mutations (ABraOM) population databases. The missense variant p.Ile657Met in the NID1 gene was the only variant that co-segregated with the disease, while absent in unaffected family members and controls. The allele frequency for this variant was <0.0001 in the gnomAD and ABbraOM databases. In silico analysis predicted the variant to be deleterious or likely damaging to the protein function. Somatic mutations in NID1 gene were found in nearly 500 cases of different cancer subtypes in the intOGen platform. Immunohistochemistry analysis showed NID1 expression in PTC cells, while it was absent in normal thyroid tissue. Our findings were corroborated using data from the TCGA cohort. Moreover, higher expression of NID1 was associated with higher likelihood of relapse after treatment and N1b disease in PTCs from the TCGA cohort. Although replication studies are needed to better understand the role of this variant in the FNMTCT susceptibility, the NID1 variant (c.1971T>G) identified in this study fulfills several criteria that suggest it as a new FNMTCT predisposing gene.

Key Words

- NID1
- FNMTCT
- papillary thyroid carcinomas
- whole-exome sequencing
Introduction

Thyroid cancers arising from follicular cells are referred to as non-medullary thyroid cancer (NMTC), while those arising from parafollicular cells are referred to as medullary thyroid cancer (MTC). Although the majority (~95%) of thyroid cancers are sporadic, a familial form has been described for both NMTC and MTC.

The familial non-medullary thyroid carcinoma (FNMT) was further divided into two groups: syndromic and non-syndromic (1). In the syndromic group, which comprises Carney complex, Werner’s syndrome, phosphatase and tensin homologue gene (PTEN) hamartoma tumour syndrome (PHTS) and familial adenomatous polyposis (FAP), NMTC occurs as a minor component of familial cancer syndromes (2, 3). As the predisposing genes for the syndromic cases have been characterized, at-risk patients can be readily identified based on genetic test.

In the non-syndromic form, thyroid cancer is the primary feature of the disease and the predisposing genes are not well defined. Even if the genetic basis of non-syndromic FNMT has not been well established, the thyroid gland shows the highest familial relative risk compared to many other well-characterized cancers with Mendelian pattern of inheritance such as breast, colon and prostate. In fact, several studies demonstrated an eight- to ten-fold increased risk of developing NMTC in first-degree relative of affected individuals than the risk observed in the general population (2, 4, 5, 6, 7, 8).

Consequently, the diagnosis of non-syndromic FNMT is primarily based on family history and it is currently defined as the occurrence of the thyroid cancer of follicular cell origin in two or more first-degree relatives in absence of other extra-thyroidal clinical manifestations (9, 10, 11, 12). Nevertheless, this classification is under debate as families with only two first-degree relatives affected by thyroid cancer may represent a casual association of the disease (13). In fact, it is estimated that nearly 62–69% of families with two first-degree relatives affected by thyroid cancer are sporadic. Though the presence of only two first-degree affected members could represent a sporadic disease, the diagnosis of FNMT should be considered and the age of onset and evidence of genetic anticipation are important parameters to account (2, 9, 10, 11, 14). Nevertheless, when three or more first-degree relatives are affected by thyroid cancer, the risk of including sporadic cases into genetic evaluation of FNMT is less than 6% (9, 13, 14).

The analysis of different kindreds with FNMT advocates for an autosomal dominant inheritance with incomplete penetrance and variable expressivity (1, 4, 5, 6, 7, 8). In the past years, several studies reported FNMT-associated chromosomal loci (3, 15, 16, 17, 18) and predisposing risk variants in over 100 genes, including SRGAP1, CHEK2, SRRM2, TIF1/NKX2, FOXE1, NOP53, HABP2, ANO7, CAV2, KANK1, PIK3CB, PKD1L1, PTPRF, BROX, RHBDD2, ATM, MAP2K5, EWSR1, POT1, TIAM1 and SPRY4 (1, 3, 7, 12, 15, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37). Recently, using whole-genomic sequencing, a group has identified variants in genes that were enriched in tumourigenic signalling pathways such as MAPK/ERK and PI3K/AKT in families with NMTC (19).

As FNMT has shown a high genetic heterogeneity, the identification of the susceptibility genes or chromosomal loci for most of FNMT cases remains to be described. Additionally, as many of these loci or susceptibility genes were not further validated in subsequent studies, genetic testing is still not recommended for diagnosis, therapeutic and follow-up decisions. However, the detection of predisposing genes or susceptibility loci is highly important, since they will undoubtedly help to identify those at high risk for FNMT, establish management strategies and assist genetic counselling.

Here we report a family with four patients affected by PTC in three generations and evidence of genetic anticipation, suggesting a true non-syndromic FNMT. Whole-exome sequencing (WES) was applied to identify a germline variant associated with an increased PTC predisposition in this family. We believe that our findings will add to FNMT diagnosis, improving cancer prevention and genetic counselling protocols.

Materials and methods

Family history

We here report a five-generation Brazilian family with four family members affected by PTC in three consecutive generations in the absence of other extra-thyroidal clinical manifestations, suggesting a true non-syndromic FNMT (Fig. 1). The family, from Natal, Rio Grande do Norte state, Brazil, was referred to our institution for genetic investigation. Family history and blood samples were obtained from consenting affected and unaffected individuals. The pathology reports confirmed the diagnosis of PTC. The proband (II-2) was diagnosed with PTC at the age of 51. The proband’s mother (II-10) and proband’s uncle (II-7) were diagnosed with PTC at the age of 75 and 76,
respectively. Proband’s son (IV-2) was diagnosed with PTC at the age of 29. Therefore, earlier age at onset of disease was observed in three successive generations, suggesting genetic anticipation. Controls, healthy individuals or individuals with PTC living in the same geographic region were included in the study. Healthy individuals were selected based on clinical and standard laboratory tests. The study was approved by the ethical and research committees of the Norte-Rio-Grandense League against Cancer (approval number 860.934) with written informed consent obtained from each participant and/or their legal representative.

**DNA isolation**

Blood samples were collected from affected and unaffected family members and controls (12 healthy individuals and 33 patients affected by PTC from the same geographic region). Genomic DNA was isolated from peripheral blood using a standard phenol/chloroform method, as described previously (38). The DNA concentration was measured using a Qubit 2.0 fluorometer (Life Technologies Corp.).

**Whole-exome sequencing**

In order to identify genes associated with non-syndromic FNMTc, DNA samples were isolated from four family members affected by PTC (II-7, II-10, III-2 and IV-2), one family member diagnosed with a thyroid benign disease (IV-3) and one unaffected member (III-1). DNA samples from two unaffected family members (III-3 and V-1) were not available for WES test. Whole exome libraries were prepared using SureSelectXT Human all exons V6 capture reagent, according manufactures’ recommendations (Agilent Technologies Inc.). Libraries were sequenced on an Illumina NextSeq550 system using NextSeq 500 High output Kit (300 cycles) (Illumina Inc., San Diego, CA, USA), targeting a mean coverage of 300×. Library preparation and sequencing were performed at the Centro de Oncologia Molecular, Hospital Sírio-libanês, São Paulo, Brazil (Fig. 2).

**Next-generation sequencing analysis**

Raw reads were initially processed for quality checking and trimming using FastQC (Babraham Institute, Cambridge) and Trimmomatic, respectively. The reads were aligned to the human reference genome (GRCh37/hg19) using Burrows–Wheeler Aligner (BWA-MEM; default parameters). Duplicates were marked by Picard (v.2.17.6; default parameters). Single nucleotides variants (SNVs) and insertions or deletions (InDels) were called using Genome Analyses Toolkit (GATK) HaplotyperCaller (version v.3.7.0; Broad Institute, Boston, MA, USA) (Fig. 2).

**Visual inspection of previously described FNMTc-associated variants**

Over the past decade many studies have identified candidate genes related to non-syndrome FNMTc. As a first step, PubMed database (‘familial non-medullary thyroid cancer’, ‘FNMTc and predisposing genes’, ‘hereditary and familial thyroid tumors’) was used to create a comprehensive list of FNMTc predisposing genes. Next, variant call format (VCF) files were manually inspected to identify variants in genes known to be associated with FNMTc (Fig. 2).
Whole Exome Analysis  
6 family members

Illumina NextSeq 550  
300x mean coverage

FASTQC  
FASTQ files  
Quality Control: Checking and Trimming

Alignment to a Reference Sequence  
GRCh37/hg19  
BWA-MEM

Variant Calling  
(GATK HaplotypeCaller)

PubMed Database Search  
Visual Inspection of Variants  
114 FNMTc predisposing Genes

PhenoDB  
Variant Annotation  
gnomAD  
MAF <1%  
Read depth > 10  
quality < 20

Filtering with Brazilian Population Databases  
(filtering out > 1%)  
ABraOM

Identification of Candidate Disease Susceptibility Variants  
(14 missense variants and 1 Indel)

Segregation Analysis  
Sanger Sequencing in Family Members (n=7) and Controls (n=45)

In silico Analysis  
Functional Prediction using VEP interface  
(SIFT, PolyPhen2, LoFtool, CADD, Condel)  
Structural Effect (HOPE) and Evolutionary conservation Analysis  
(Uniprot, Jalview)

Immunohistochemistry  
Analysis of FFPE thyroid sections

Data mining for Somatic Mutations  
TCGA, IntOGen and Cosmic

TCGA Gene Expression  
Candidate gene expression in TCGA-ThCA cohort.

Figure 2  
Flow chart showing the strategy and pipeline used in this study.

**Variant annotation and filtering**

Analysis of VCF files was performed using PhenoDB analysis tool (https://phenodb.org/) (39). PhenoDB was used to analyse frame shift, nonsense or missense variants that were shared between the four affected family members based on dominant or recessive modes of inheritance. Variants were filtered out according to read depth (<10 reads), genotype quality (<20) and frequency in the Genome Aggregation Database (gnomAD) and Online Archive of Brazilian Mutations (ABraOM). The gnomAD (https://gnomad.broadinstitute.org) spans data from 125,748 exomes and 15,708 genomes from healthy unrelated individuals from 60 populations. The ABraOM (http://abraom.ib.usp.br) database was also used to assess the frequency of the annotated variants. This database is composed of an elderly cohort of over 1500 Brazilian individuals aged 60 or older, including immigrant descendants from different continents and individuals from various Brazilian states (40). Only variants with MAF (minor allele frequency) <1% (gnomAD and ABraOM) were annotated (Fig. 2).

**Sanger sequencing analysis**

Sanger sequencing was used to validate the selected variant and to investigate its segregation in this family. Whole blood DNA from four family members affected by PTC (II-7, II-10, III-2 and IV-2) and three unaffected (III-1, IV-3 and V-1) were submitted to Sanger sequencing. Although the DNA of the 1-year-old child (V-1) was not available for WES test, her DNA was available for the validation analysis. The selected variant was also investigated in 12 healthy individuals and 33 patients affected by PTC from the same geographic region. Primers were designed to amplify the target region of NID1 gene using Primer3 v.0.4.0 (http://primer3.ut.ee) (F: CAATGTCAAAGAGTGGGGTTT; R: GCAGCTCTCGGTGGACAG). Amplified PCR product was purified and sequenced using the BigDye Terminator (Life Technologies) in the ABI 3100 Genetic Analyzer (Applied Biosystems).

**In silico analysis**

The effect of the amino acid substitution on NID1 protein structure and function was performed using various available plugins via the Ensembl Variant Effect Predictor interface (ensembl.org/info/docs/tools/vep) such as Sorting Intolerant From Tolerant (SIFT, http://sift.jcvi.org), Polymorphism Phenotyping v2 (Polyphen-2, http://genetics.bwh.harvard.edu/pph2), Loss-of-function (LoFtool), Combined Annotation-Dependent Depletion (CADD, https://cadd.gs.washington.edu), phred score (which is normalized to all potential SNVs providing an externally comparable unit for analysis) and Consensus Deleteriousness (Condel) scores.

The online software HOPE (http://www.cmbi.ru.nl/hope/method) was used to analyse the structural effect of the amino acid change in the protein. PsiPred (http://bioinf.cs.ucl.ac.uk/psipred) was used to predict amino acid changes in the protein secondary structure.

Uniprot database (https://www.uniprot.org/) and Jalview software (version 2.11.1.3) were used to analyse the evolutionary conservation of amino acids by calculating
the conservation score. This software uses a calculation based on AMAS (analysis of multiply aligned sequences) to determine the amino acid conservation score. To obtain the value, ten amino acid properties were evaluated to create an index that ranges from 1 to 11 (hydrophobic, polar, small, proline, tiny, aliphatic, aromatic, positive, negative and charged). A maximum score (score 11) is represented by ‘+’; if the entire properties match but the amino acid is different across species, the signal ‘+’ is used.

Immunohistochemistry analysis of available PTC samples

Formalin-fixed, paraffin embedded (FFPE) thyroid samples from two family members affected by PTC (II-7 and IV-2) were available for immunohistochemistry (IHC) analysis. Tissue slides (4 µm) were deparaffinized with xylene and hydrated through a graded alcohol series. Endogenous peroxidase activity was blocked by placing slides in 15% hydrogen peroxide solution for 10 min. Antigen retrieval was performed in Tris/EDTA buffer pH 9.0 (Envision Flex Target Retrieval Solution, Dako Agilent) in a steamer for 20 min. After cooling to room temperature for 1 h, nonspecific binding sites were blocked with 5% BSA in PBS. The slides were washed twice with 1x PBS and incubated with mouse MAB against NID1 (R&D Systems, Minneapolis, MN, USA; Cat no. MAB2570). The samples were incubated with NID1 antibody (1:50) overnight at 4°C, followed by incubation with secondary antibody-polymer horseradish peroxidase (HRP) for 60 min (VisUCyte™Mouse, Cat no. VCTS001). Slides were stained with DAB chromogen solution (VisUCyte™ Mouse). The nuclei was counterstained with haematoxylin and analysed using a light microscope.

NID1 analysis in sporadic cancers

The prevalence of the selected variant was assessed in nearly 500 sporadic cases of PTC from thyroid cancer database of The Cancer Genome Atlas (TCGA; https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga). Additionally, the incidence of the selected variant was also assessed in 1,500,000 human cancers present in the Catalogue of Somatic Mutations in Cancer database (COSMIC v92; https://cancer.sanger.ac.uk/cosmic) and in the Integrative OncoGenomics analysis (intOGen; https://www.intogen.org/search). The intOGen evaluated the mutational status in 28,076 samples of 66 cancer types, identified cancer genes and pinpointed their putative mechanism of action across cancer types. The IntOGen comprised 684 thyroid cancers from four different cohorts, including 488 PTC cases from TCGA cohort.

The expression of NID1 was also evaluated in PTCs from TCGA cohort, and its expression was correlated with available clinical and pathological data. DESeq2 was used to test for differential expression (41). For statistical analysis we used Kruskal–Wallis test followed by Dunn post hoc test.

Results

Screening of variants in candidate predisposing genes for non-syndromic FNMTC

To identify genetic variants associated with FNMTC, WES analysis was performed in six members (four affected and two unaffected) from a five-generation Brazilian family with PTC and a Mendelian-like inheritance (Fig. 1). We obtained on average 57 million reads per sample (49–59 million reads). After filtering, the reads were mapped to the human reference genome (hg19) and variant calling was performed using GATK programme (Fig. 2).

As first step, we mined the literature for candidate non-syndromic FNMTC susceptibility genes. In total, we identified 114 candidate genes across 32 independent studies (15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 31, 33, 34, 35, 36, 37, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60). A list of included genes and studies is provided in Supplementary data (Supplementary Table 1, see section on supplementary materials given at the end of this article). Next, WES data from affected and unaffected family members were manually inspected for variants in all candidate genes. No germline variant in the protein-coding regions of 114 pre-selected FNMTC susceptibility genes was found to co-segregate with the disease phenotype in this family, suggesting that a novel candidate variant could be implicated in the genetic predisposition to FNMTC.

The p.Ile657Met NID1 variant co-segregates with the phenotype in a five-generation family

Novel susceptibility variants for FNMTC were inspected using PhenoDB, a robust database that enables efficient and adjustable sequence analysis coupled to the phenotypic and other clinical information. We identified 15 missense variants located in 13 genes in all affected family members (Table 1). To identify if any biological pathway is enriched among the 13 candidate genes, the list of genes was used
as input for enrichment (maayanlab.cloud/Enrichr). Enrichment analysis identified that three genes (ZNF717, ZNF578 and ZNF667) were mapped to Herpes simplex virus 1 infection KEGG pathway (P-adj < 0.05).

However, 14 out of 15 missense variants were also reported in healthy controls from the same geographic region. The c.1971T>G heterozygous missense variant (NM_002508) in the exon 8 of the Nidogen-1 (NID1) gene, which resulted in an isoleucine (ATI) to methionine (ATG) substitution at residue 657, is the unique variant that co-segregated with PTC phenotype in this family (Fig. 3).

**Sanger analysis confirmed the presence of NID1 variant in affected family members**

Sanger analysis confirmed the presence of the p.Ile657Met NID1 variant in all family members affected with PTC, and it was absent in unaffected family members (III-1 and IV-3) (Fig. 3B). Validation analysis revealed that the 1-year-old child (V-1) was a carrier of the p.Ile657Met variant, while it was absent in healthy controls and individuals affected by PTC from the same geographic region (Fig. 3C).

Longer follow-up will help determine the pathogenicity of the p.Ile657Met NID1 genetic variant in this family, particularly in the 1-year-old unaffected child who is a carrier of the NID1 variant.

**Filtering with gnomAD and ABraOM population databases**

The gnomAD data set (v2.1.1) showed that the NID1 p.Ile657Met variant is extremely rare as it was found at a MAF of 0.00007798. The gnomAD v3.1 (non-cancer) spans 74,023 whole-genome sequences from unrelated individuals from over 60 distinct populations, who were not ascertained for having cancer in a cancer study (Table 2). The p.Ile657Met variant was reported in gnomAD v3.1 (non-cancer) also at exceptionally low allele frequency (0.00009476). Although the gnomAD v3.1 covers the data from 1000 Genomes (1KG), we examined it independently as it lists human genetic variations in 3202 normal, healthy individuals from 26 different populations from many different locations around the globe (61). The NID1 p.Ile657Met variant was not reported in 1KG.

We also assessed the ABraOM repository of variant frequencies, a census-based cohort of over 1500 elderly individuals from São Paulo, Brazil (40). Notably, Brazilian population is constituted by nearly 500 years of admixture between Africans, Europeans and native Americans (Table 2). The p.Ile657Met NID variant was found at very low prevalence (<0.01%) and, therefore, considered as an extremely rare variant in the Brazilian population.

**Pathogenicity prediction scores**

To predict the likely functional impact of the missense p.Ile657Met variation on NID1 protein stability, activity or folding, we used computational prediction methods (Table 1). Interestingly, most of the prediction scores presumed that the amino acid change, which is located at a well-conserved position at the NID1 protein, is deleterious (SIFT, Condel) or probably damaging (PolyPhen-2, LoFtoll to protein function (Table 1).

The HOPE analysis showed that the original WT residue and newly introduced mutant residue often differ in size, charge and hydrophobicity value properties. Not only the methionine residue is bigger than the isoleucine but also has different properties (Fig. 4A and B). The residue is buried in the core of a domain within the protein, and the differences between the WT and mutant residue are likely to disturb the core structure of this domain and alter its function (Fig. 4C and D).

**Evolutionary conservation analysis**

The Uniprot database and Jalview software (version 2.11.1.3) were used to compare the NID1 residues conservation across species. Amino acid sequences of NID1 or the orthologous protein were obtained for the following species: *Homo sapiens* (Uniprot accession P14543), *Mus musculus* (Uniprot accession P10493), *Caenorhabditis elegans* (Uniprot accession Q93791), *Pan troglodytes* (Uniprot accession H2Q1E0), *Macaca mulatta* (UniProt accession F7DF16), *Canis lupus familiaris* (UniProt accession A0ASF4C1R1), *Bos taurus* (UniProt accession A6QNY6), *Rattus norvegicus* (UniProt accession F1LM84), *Felis catus* (UniProt accession M3WED3), *Sus scrofa* (UniProt accession F1RGY5) and *Equus caballus* (UniProt accession F6XH19).

The conservation, quality and consensus tracks from Jalview were calculated for each column showed in the histogram. As observed for selected regions of the alignments, the annotation scores (ranging from 1 to 11) reflect conservation of the physicochemical properties in the particular column of the alignment. The absolutely conserved residues are marked with a yellow asterisk. Other amino acids are highlighted with different colours, according to the conservation scores in the consensus sequence. Less conserved positions are shown in darker colours with decreasing score. The residue I (Ile) at position...
Table 1  Candidate susceptibility variants identified in this study.

| Gene name | Position | Nucleotide change | Aminoacid change | dbsNP | Mutation classification | SIFT | PolyPhen2 | LoFtool | CADD phred | Condel |
|-----------|----------|-------------------|------------------|-------|-------------------------|------|-----------|---------|------------|--------|
| CAPN9     | chr1:230907829 | G>A               | p.G224R          | rs143186556 | Nonsynonymous SNV       | Deleterious (0) | Probably damaging (1) | -       | 29.7       | Deleterious (0.945) |
| NID1      | chr1:236189209 | T>G               | p.I657M          | rs199655594 | Nonsynonymous SNV       | Deleterious (0.02) | Probably damaging (0.98) | -       | 12.35      | Deleterious (0.807)  |
| TRANK1    | chr3:36887865 | T>C               | p.T1645A         | -      | Nonsynonymous SNV       | Deleterious (0) | Probably damaging (0.95) | -       | 25.0       | Deleterious (0.902)  |
| POMGNT2   | chr3:43121810 | C>T               | p.D372N          | rs147129872 | Nonsynonymous SNV       | Tolerated (0.39) | Benign (0.159)      | -       | -          | 23.2   |
| ZNF717    | chr3:75786355 | T>A               | p.H63L           | rs201105907 | Nonsynonymous SNV       | Tolerated (0.67) | Benign (0.121)      | -       | -          | 17.65  |
| DGKQ      | chr4:961091  | C>T               | p.R349Q          | rs183334682 | Nonsynonymous SNV       | Tolerated (0.58) | Benign (0)           | -       | -          | 5.565  |
| MUC5B     | chr1:1260217 | C>A               | p.D1138E         | rs79638064 | Nonsynonymous SNV       | Tolerated (0.13) | Benign (0.229)      | -       | 2.651      | -      |
| MUC5B     | chr1:1260219 | C>T               | p.A1139W         | rs75935363 | Nonsynonymous SNV       | Deleterious (0.05) | Probably damaging (0.93) | -       | -          | -      |
| CACNA1H   | chr1:6:1265318 | G>A              | p.A1700T         | rs761499667 | Nonsynonymous SNV       | Deleterious (0.03) | Probably damaging (0.934) | -       | 17.78      | Deleterious (0.722)  |
| C16orf96  | chr1:6:4644482 | G>A              | p.G1000R         | rs142136567 | Nonsynonymous SNV       | Deleterious (0.03) | Possibly damaging (0.851) | -       | 23.3       | Deleterious (0.744)  |
| IL4I1     | chr1:9:50399161 | G>A             | p.R55W           | rs140225154 | Nonsynonymous SNV       | Tolerated (0.05) | Possibly damaging (0.727) | -       | 25.0       | Deleterious (0.692)  |
| ZNF578    | chr1:9:53015322 | G>A              | p.G563E          | rs566285018 | Nonsynonymous SNV       | Deleterious (0.03) | Possibly damaging (0.627) | -       | 23.4       | Deleterious (0.613)  |
| ZNF667    | chr1:9:56953511 | T>C             | p.K285E          | rs61740676 | Nonsynonymous SNV       | Tolerated (0.79) | Benign (0)           | -       | 21.2       | Deleterious (0.583)  |
| CDTS2P2   | chr1:2:58217744 | >TC              | p.R211fs         | -      | Frameshift insertion    | -    | -         | -       | 0.009       | -      |
| DGKQ      | chr4:961091  | C>T               | p.R349Q          | rs183334682 | Nonsynonymous SNV       | Tolerated (0.58) | Benign (0)           | -       | -          | -      |

SIFT score ranges from 0.0 (deleterious) to 1.0 (tolerated). Variants with score from 0.0 to 0.05 are considered deleterious. PolyPhen-2 score ranges from 0.0 (tolerated) to 1.0 (deleterious), variants with score from 0.85 to 1 are more confidently predicted to be damaging. LoFtool provides a rank of genic intolerance and consequent susceptibility to disease based on the ratio of loss-of-function; the lower the LoFtool gene score percentile the most intolerant is the gene to functional variation. CADD phred score of 20 means that a variant is among the top 1% of deleterious variants in the human genome. Condel calculates consensus deleteriousness score of missense mutations; variants with score above 0.49 are considered deleterious.
657 displayed a score of 9 across the different species, which means a conservation of nine among ten amino acid properties is considered to create this index (Fig. 4E). It is well known that residues more relevant for functioning of the protein are more evolutionary conserved (i.e. preserved). As the evolutionary conservation of the corresponding site correlates with the effect of mutation, this result suggested that the mutation found in this residue of NID1 protein likely impacts the function of the protein.

Figure 3
(A) The graphic shows the p.Ile675Met variant in the NID1 gene and its localization relative to the respective protein domains. (B) Sanger sequencing confirmed the c.1971T>G substitution highlighted in yellow (antisense strand), located within exon 8 of the NID1 gene in all individuals affected by PTC (II-7, II-10, III-2 and IV-2) and its absence in an unaffected family member (III-1) and in the patient with benign thyroid lesion (IV-3). The 1-year-old child (V-1) is a carrier for the c.1971T>G substitution. (C) Sanger sequencing analysis confirming the absence of NID1 c.1971T>G variant in unaffected individuals from the same geographic region.

| Table 2 | The p.Ile657Met allele frequency in different populations. |
|---------|-------------------------------------------------------------|
|         | ABraOM SABE WES (n = 609) | AbraOM SABE-WGS (n = 1171) | gnomAD (non-cancer) v3.1* (n = 74,023) | gnomAD v2.1.1** (n = 141,456) | 1KGP (n = 3202) |
| European (non-Finnish) | ND | ND | 0.0001699 | 0.0001478 | ND |
| Latino | ND | ND | 0.0001327 | 0.00005645 | ND |
| South Asian | ND | ND | 0.0000000 | 0.00000000 | 0.000 |
| African | ND | ND | 0.00002436 | 0.00000000 | 0.000 |
| Ashkenazi Jewish | ND | ND | 0.000 | 0.000 | ND |
| East Asian | ND | ND | 0.000 | 0.000 | 0.000 |
| European (Finnish) | ND | ND | 0.000 | 0.000 | 0.000 |
| Middle Eastern | ND | ND | 0.000 | 0.000 | ND |
| Brazilian | 0.000821 | 0.000427 | ND | ND | ND |
| Male | 0.0000821 | 0.000427 | 0.00005576 | 0.00009152 | 0.000 |
| Female | 0.0001316 | 0.000427 | 0.00001316 | 0.00006194 | 0.000 |
| Allele frequency | 0.000821 | 0.000427 | 0.00009476 | 0.00007798 | 0.000 |
Considering conservation and orthologue structure of NID1 protein across species, a phylogenetic tree was constructed using neighbour joining. This methodology implemented in jalview applies a greedy algorithm to find the tree with the shortest branch lengths. As expected, NID1 human protein share higher similarity with Pan troglodytes (UniProt accession H2Q1E0) and Macaca mulatta (UniProt accession F7DF16) (Fig. 4F).

NID1 p.Ile657Met variant was absent in sporadic human cancers

The analysis of the NID1 p.Ile657Met variant in nearly 500 sporadic PTC cases from TCGA and in nearly 1,500,000 human cancers present in COSMIC showed an absence of this variant. Although the p.Ile657Met variant was absent in sporadic PTC cases, we looked for variants within NID1 gene in the IntOGen mutation platform, which collect and analyse somatic mutations in thousands of tumour genome from different cancers subtypes, including COSMIC, TCGA and International Cancer genomes Consortiums. Although no somatic mutation was in found at codon 657 of NID1 gene, the analysis disclosed 495 coding variants throughout the NID1 gene in different cancer types. Most of them (65%) are missense mutations, including the p.Thr863Ala variant in one PTC from the TCGA cohort.

NID1 expression in PTC samples

Representative haematoxylin and eosin (H&E) stained sections used to confirm the diagnosis of PTC are showed (Fig. 5A and B). IHC analysis performed in the available PTC thyroid tissue samples showed NID1 staining (brown) in the tumour cells (TC) (Fig. 5C and D), while it was negative in adjacent normal thyroid (NT) tissue (Fig. 5C).

NID1 expression in PTC from TCGA cohort was associated with more aggressive phenotype

As NID1 was expressed in PTC samples from two affected members and it was absent in normal thyroid by IHC, we next investigated the mRNA expression of NID1 in PTC from TCGA cohort and correlated with available clinical and pathological parameters. NID1 expression was higher in PTC and paired lymph node metastases samples compared to its expression in normal thyroid tissues (P-value and P-adj. < 0.05) (Fig. 6A).

Moreover, NID1 expression was higher in PTC samples from patients who showed microscopic (R1) and R2 (macroscopic) residual tumours than the expression observed in PTC samples from patients who showed no residual disease present or complete remission (R0) and patients in whom the presence of residual cancer could not be assessed (RX) (P-value < 0.05) (Fig. 6B). Therefore, higher expression of NID1 was found in PTC with higher likelihood of relapse after treatment. We subsequently correlated its expression with lymph node (N) stage. No significant difference was observed in the expression of NID1 in patients with no evidence of regional lymph node metastasis (N0) and those with lymph node disease (N1). When N1 group was further divided into N1a (central node metastasis only) and lateral neck lymph nodes (N1b), higher expression of NID1 was detected in patients in the N1b group than its expression in patients in the group N0 (P-value and P-adj. < 0.05) (Fig. 6C). No difference was observed in other available clinical pathological features.

Remarkably, NID1 amplification was found in 4/496 (0,81%) PTC cases from the TCGA cohort. All together, these results not only supported our findings but also suggested that NID1 expression was associated with a more aggressive and diffuse PTC.

Discussion

Despite the advances in understanding the genetic basis of non-syndromic FNMTC, no susceptibility genes have been identified. Therefore, no genetic test is currently available to identify at-risk patients and the diagnosis is based on the family history.

We herein report a five-generation Brazilian family with positive family history of PTC generations in the absence of other extra-thyroidal clinical manifestations, suggesting a true non-syndromic FNMTC. Regarding disease presentation, we observed the occurrence of the disease in three consecutive generations in both sexes, which is consistent with an autosomal dominant inheritance pattern. Although no significant difference was observed in the age of onset of the affected family member in the second generation (IV-2 and IV-3), individuals in third and fourth generations manifested the disease at an earlier age. The observed decreased age at onset in successive generations suggested genetic anticipation and reinforced the premise that is a true familial disease rather than fortuitous association.

https://ec.bioscientifica.com
https://doi.org/10.1530/EC-21-0406
© 2022 The authors
Published by Bioscientifica Ltd

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
Modelling of NID1 variant on 3D structure of the protein. (A) The 3D folding shows the isoleucine to methionine change (magenta) at position 657 in the NID1 protein (grey). The mutant amino acid differs in its hydrophobic and charge proprieties. The residue in the reference protein (WT) is smaller than the mutant residue (http://www.cmbi.ru.nl/hope). (B) The protein structure (grey) was zoomed in to show the WT residue (green) and the mutant residue (red). (C and D) Secondary structure map of the WT and mutant NID1 protein, respectively. Magenta represents α-helix structures and yellow represents β-sheets. (E) Amino acid sequence alignment of NID1 with ten orthologous proteins. The conserved isoleucine residue at position 657 is highlighted. Conservation scores is 9, which means that amino acids in this positions match in nine of ten properties across species. A maximum score is represented by '*', the signal '+' indicates that the entire properties match but the amino acid is different across species. (F) Phylogenetic tree of 11 sequences of NID1 is depicted. This tree was constructed based on amino acids alignments. NID1 human protein share stronger sequence similarity with Pan troglodytes (UniProt accession H2Q1E0) and Macaca mulatta (UniProt accession F7DF16).
To identify the underlying genetic basis of the disease in this family, WES was performed. As part of the strategy, affected and unaffected family members were initially screened for variants in 114 genes previously described as associated with non-syndromic FNMTC. As variants in the candidate genes were excluded, we next explored the WES data for novel candidate genes for non-syndromic FNMTC. We identified a novel missense variant within \textit{NID1} gene co-segregating with the disease in this family. Sanger sequencing confirmed that the \textit{NID1} p.Ile675Met variant co-segregated with the disease phenotype. Remarkably, the only unaffected carrier was the 1-year-old child. As penetrance is age dependent, longer follow-up is needed to better define the penetrance of PTC in this carrier.

\textit{NID1} gene is located at 1q42.3 and encodes a member of the nidogen family of basement membrane glycoproteins. The protein interacts with several other components of basement membranes and may play a role in cell interactions with the extracellular matrix. Gene Ontology (GO) annotations related to this gene include calcium ion binding and laminin binding.

Although several lines of evidence support that \textit{NID1} variant is a new candidate variant associated with FNMTC, this study has three main limitations. First, \textit{NID1} variant was not identified in other families with non-syndromic FNMTC. However, it is important to mention that using WES and comprehensive pathway enrichment analysis, \textit{NID1} was previously reported as one of the mutated genes associated with extracellular matrix organization pathway in a family with PTC, breast cancer and benign lesions in Brazil. The authors suggested that this gene might have pathogenic effects in the risk of familial PTC development (53). Secondly, although all affected member of this family underwent WES, we did not test some unaffected individuals, as DNA was not available. Thirdly, no functional analysis was performed for this specific variant found in exon 8 of \textit{NID1} gene. Although functional analysis would help to define the role of \textit{NID1} as predisposing gene for FNMTC, it is important to notice that though functional studies can help supporting pathogenicity, not all analysis are effective in predicting the impact of a variant in a protein function.

In an attempt to overcome these limitations, several parameters commonly used to classify variants were considered such as MAF, \textit{in silico} tools analysis for interpretation of pathogenicity of the variant (functional prediction, structural effect and evolutionary conservation analysis) and \textit{NID1} expression in thyroid tumour samples from affected members and in PTC samples from TCGA cohort. According to the parameters used to categorize variants by frequency, the p.Ile675Met \textit{NID1} variant is extremely rare in control populations from gnomAD data set (MAF < 0.0001) and from elderly individuals from São Paulo, Brazil (MAF < 0.001) and it is absent in the 1000 Genomes project.

\textit{In silico} analysis is a rapid and a reliable way to improve our understanding of how an amino acid substitution could affect the protein structure and function. Using different platforms, we attempted to cover the major

Figure 5
Haemotoxylin and eosin staining of two affected family members II-7 (A) and IV-2 (B) confirmed the diagnosis of papillary thyroid cancer. \textit{NID1} expression analysis in thyroid sections of two family members affected with PTC. Staining of tumour cells (TC) was observed in thyroid sections of cases II-7 (C) IV-2 (D), while adjacent thyroid normal (NT) cells showed negative staining (C).
aspects influencing protein function. The variant was classified by in silico tools most commonly used in clinical laboratories as probably pathogenic (PolyPhen-2 and LoFtool) or deleterious (SIFT and Condel) to the NID1 protein function or structure. Although the impact of an amino acid substitution on protein function differs across predictor methods, as they are based on different criteria such as the evolutionary conservation of the amino acid, the impact of the context and the biochemical consequence of the amino acid substitution were rather similar. Additionally, amino acid alignment confirmed that the residue Ile at position 675 of the NID1 protein is conserved across multiple species (score 9). Amino acid substitution at conserved positions more likely leads to deleterious effects on gene functions.

In summary, these data combined suggest that the missense variant likely has an effect on protein function and it is potentially associated with phenotype in this family.
Expression analysis performed in thyroid sections obtained from affected family members suggests that \textit{NID1} might play a role in the pathogenesis of PTC, as it is expressed in thyroid cancer cells from PTC samples but is not expressed in normal thyroid follicular cells. Moreover, the results of PTC from TCGA cohort showed that compared with normal thyroid tissues, the expression of NID1 was higher in cancer cells. Higher expression of NID1 was associated with higher likelihood of relapse after treatment and N1b disease, suggesting that NID1 expression is associated with a more aggressive and diffuse PTC.

Data from literature also recognized the role of NID1 in the pathogenesis and progression of different cancer subtypes \cite{62, 63, 64}. As example, it was showed that high expression of NID1 significantly correlated with shorter overall survival of ovarian cancer patients. The authors demonstrated that the ectopic expression of NID1 in an adenocarcinoma cell line promoted cell migration, invasion and chemoresistance and that the underlying mechanism involves, at least in part, the activation of ERK/MAPK signalling pathway \cite{62}. \textit{In vitro} analysis further revealed multiple prometastatic functions of NID1, including enhanced cell migration and invasion, and improved vascular tube formation capacity. Importantly, higher expression of NID1 was associated with an increased risk of developing distant metastasis and poor clinical outcomes. The authors suggested that NID1 is a prometastatic-secreted protein that may develop as a new biomarker for disease progression and therapeutic target in breast cancer and melanoma \cite{65}. Higher levels of NID1 expression was also correlated with the advanced stages of oral cavity squamous cell carcinoma (OSCC) and facilitated gastric cancer cell migration and invasion \textit{in vitro} and poor survival outcome in gastric cancer patients, highlighting the role of NID1 in gastric cancer progression \cite{66}. Interestingly, NID1 signalling pathway was also associated with the pathogenesis of Hirschsprung disease (HSCR) \cite{67}. Recently, it was demonstrated that the loss of \textit{NDRG4} (N-MYC downstream-regulated gene 4) promotes colorectal cancer via increased release of \textit{NID1} and \textit{FBLN2} \cite{68}.

Recently, the thyroglobulin interactome map showed that NID1 is one of the proteins that comprise the secretory proteostasis network associated with thyroglobulin mutation-specific changes that impair proper production, folding or processing leading to dyshormonogenesis \cite{69}. However, further analysis is needed to better understand the role of \textit{NID1} in thyroid cancer process.

In conclusion, we report individuals with a family history of PTC and evidence of clinical anticipation in successive generations, suggesting a true non-syndromic FNMT. Although the affected members were negative for variants in genes previously reported as associated with FNMT, WES analysis identified a novel p.I657M variant within \textit{NID1} gene that segregates with the phenotype and is absent in controls from same geographic region. The p.Ile657Met allele frequency is extremely low in control populations, and multiple lines of computational evidence support deleterious/damaging effect of the amino acid substitution on NID1 protein. Although further analysis will help to better understand the role of this variant in the pathogenesis of FNMT, \textit{NID1} variant fulfills several criteria that suggest it as a new candidate FNMT predisposing gene.

**Supplementary materials**

This is linked to the online version of the paper at \url{https://doi.org/10.1530/EC-21-0406}.

**Declaration of interest**

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

**Funding**

This work was supported by grants from the São Paulo Research Foundation (FAPESP 2014/06570-6 and 2018/23497-1). P A F G is funded by FAPESP (grant 2018/15579-8). J M C and P A F G are investigators of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (grant 304534/2018-8 for J M C).

**Acknowledgements**

The authors thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for their financial support and the patients diagnosed with cancer attended to at the outpatient clinic of a referral hospital (Norte-Rio-Grandense League Against Cancer).

**References**

1. Wang Y, Liyanarachchi S, Miller KE, Nieminen TT, Comiskey DE, Li W, Brock P, Symer DE, Akagi K, DeLap KE, \textit{et al.}. Identification of rare variants predisposing to thyroid cancer. \textit{Thyroid} 2019 \textbf{29} 946–955. \url{https://doi.org/10.1089/thy.2018.0736}
2. Houlston RS & Stratton MR. Genetics of non-medullary thyroid cancer. \textit{QJM} 1995 \textbf{88} 685–693.
3. Peiling Yang S & Ngeow J. Familial non-medullary thyroid cancer: unraveling the genetic maze. \textit{Endocrine-Related Cancer} 2016 \textbf{23} R577–R595. \url{https://doi.org/10.1530/ERC-16-0067}
4. Moses W, Weng J & Kebebew E. Prevalence, clinicopathologic features, and somatic genetic mutation profile in familial versus sporadic nonmedullary thyroid cancer. \textit{Thyroid} 2011 \textbf{21} 367–371. \url{https://doi.org/10.1089/thy.2010.0256}
5. Bauer AJ. Clinical behavior and genetics of nonsyndromic, familial nonmedullary thyroid cancer. \textit{Frontiers of Hormone Research} \textbf{41} 2013 141–148. \url{https://doi.org/10.1159/000345674}
6 Frank C, Fallah M, Sundquist J, Hemminki A & Hemminki K. Population landscape of familial cancer. *Scientific Reports* 2015 5 12891. (https://doi.org/10.1038/srep12891)

7 Ammar SA, Alobua WM & Kebebew E. An update on familial nonmedullary thyroid cancer. *Endocrine* 2020 68 502–507. (https://doi.org/10.1007/s12020-020-02250-3)

8 Malchoff CD & Malchoff DM. Familial nonmedullary thyroid carcinoma. *Cancer Control* 2006 13 106–110. (https://doi.org/10.1665/10732480(2006)13[106:E]2.0.CO;2)

9 Capezzone M, Robenshtok E, Cantara S & Castagna MG. Familial nonmedullary thyroid cancer: a critical review. *Journal of Endocrinological Investigation* 2021 44 943–950. (https://doi.org/10.1007/s00050-020-01435-x)

10 Nixon IJ, Suárez C, Simo K, Sanabria A, Ángelos P, Rinaldo A, Rodrigo JP, Kowalski LJ, Hartl DM, Himni ML, et al. The impact of family history on non-medullary thyroid cancer. *European Journal of Surgical Oncology* 2016 42 1455–1463. (https://doi.org/10.1016/j.ejso.2016.08.006)

11 Klubo-Gwiazdowska J, Yang L, Mertel R, Patel D, Nihoul N, Merino MJ, Skarulis M, Sadowski SM & Kebebew E. Results of screening in familial non-medullary thyroid cancer. *Thyroid* 2017 27 1017–1024. (https://doi.org/10.1089/thy.2016.0668)

12 Diguigiovanni C & Bonora E. Genetics of familial non-medullary thyroid carcinoma (FNMC). *Cancers* 2021 13 2178. (https://doi.org/10.3390/cancers13092178)

13 Charles ND. On the prevalence of familial nonmedullary thyroid cancer in multiply affected kindreds. *Thyroid* 2006 16 181–186. (https://doi.org/10.1089/thy.2006.16.181)

14 Capezzone M, Fralassi N, Secchi C, Cantara S, Brilli L, Pilli T, Maino F, Forleo R, Pacini F, Cevenini G, et al. Long-term clinical outcome in familial and sporadic papillary thyroid carcinoma. *European Thyroid Journal* 2020 9 213–220. (https://doi.org/10.1159/000506955)

15 Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Berghorsson JT, He H, Blondal T, Geller J, Jakobsdottir M, et al. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nature Genetics* 2009 41 460–464. (https://doi.org/10.1038/ng.339)

16 McKay JD, Lesueur F, Jonard I, Pastore A, Williamson J, Hoffman L, Burgess J, Duffett A, Papotti M, Stark M, et al. Localization of a susceptibility gene for familial nonmedullary thyroid carcinoma to chromosome 2q21. *American Journal of Human Genetics* 2001 69 440–446. (https://doi.org/10.1086/321979)

17 Malchoff CD, Sarafraz M, Tendler B, Fugazzola L, Arnold A & Malchoff DM. Papillary thyroid carcinoma associated with papillary renal neoplasia: genetic linkage analysis of a distinct heritable tumor syndrome. *Journal of Clinical Endocrinology and Metabolism* 2000 85 1758–1764. (https://doi.org/10.1210/jcem.85.5.6557)

18 Cavaco BM, Batista PF, Sobrinho LG & Leite V. Mapping a new familial thyroid epithelial neoplasia susceptibility locus to chromosome 8p23.1-p22 by high-density single-nucleotide polymorphism genome-wide linkage analysis. *Journal of Clinical Endocrinology and Metabolism* 2008 93 4426–4430. (https://doi.org/10.1210/jc.2008-0449)

19 Srivastava A, Kumar A, Giangiobbe S, Bonora E, Hemminki K, Forsti A & Bandapalli OR. Whole genome sequencing of familial non-medullary thyroid cancer. *American Journal of Human Genetics* 2019 105 10566. (https://doi.org/10.1016/j.ajhg.2019.10566)

20 Pereira JS, Silva JG da, Tomaz RA, Pinto AE, Bugalho MJ, Leite V & Cavaco BM. Identification of a novel germline FOXE1 variant in patients with familial non-medullary thyroid carcinoma (FNMC). *Endocrine* 2015 49 204–214. (https://doi.org/10.1007/s12020-014-0470-0)

21 Tomisie J, He H, Akagi K, Liyanarachchi S, Pan Q, Bertani B, Nagy R, Symer DE, Blencowe BJ & de la Chapelle A. A germline mutation in SRM2M, a splicing factor gene, is implicated in papillary thyroid carcinoma predisposition. *Scientific Reports* 2015 5 10566. (https://doi.org/10.1038/srep10566)

22 Gaia SK, Jia L, Merino MJ, Agarwal SK, Zhang L, Cam M, Patel D & Kebebew E. Germline HRB2 mutation causing familial nonmedullary thyroid cancer. *New England Journal of Medicine* 2019 373 448–455. (https://doi.org/10.1056/NEJMoa1502449)

23 Capezzone M, Cantara S, Marchisotta S, Filetti S, De Santi MM, Rossi B, Ronga G, Durante C & Pacini F. Short telomeres, telomerase reverse transcriptase gene amplification, and increased telomerase activity in the blood of familial papillary thyroid cancer patients. *Journal of Clinical Endocrinology and Metabolism* 2008 93 3950–3957. (https://doi.org/10.1210/jc.2008-0372)

24 Zhu J, Wu K, Lin Z, Bai S, Wu J, Li P, Xue H, Du J, Shen B, Wang H, et al. Identification of susceptibility gene mutations associated with the pathogenesis of familial nonmedullary thyroid cancer. *Molecular Genetics and Genomic Medicine* 2019 7 e1015. (https://doi.org/10.1002/mggn.10315)

25 Orosi A, Gara SK, Mora M, Halperin I, Martínez S, Alfayate R, Kebebew E & Oriola J. NOP53 as a candidate modifier locus for familial non-medullary thyroid cancer. *Genes* 2019 10 899. (https://doi.org/10.3390/genes10110899)

26 Landi I, Ruiz-Llorente S, Montero-Conde C, Ingład-Pérez I, Schiavi F, Leskelä S, Pita G, Milne R, Maravall J, Ramos I, et al. The variant rs1867277 in FOXE1 gene confers thyroid cancer susceptibility through the recruitment of USF1/USF2 transcription factors. *PLoS Genetics* 2009 5 e1000637. (https://doi.org/10.1371/journal.pgen.1000637)

27 Mancikova V, Cruz B, Ingład-Pérez L, Fernández-Rozadilla C, Landi I, Cameselle-Teijeiro J, Celeiro C, Pastor S, Velázquez A, Marcos R, et al. Thyroid cancer GWAS identifies 10q26.12 and 6q14.1 as novel susceptibility loci and reveals genetic heterogeneity among populations. *International Journal of Cancer* 2015 137 1870–1878. (https://doi.org/10.1002/ijc.29557)

28 Cirello V, Colombo C, Persani L & Fugazzola L. Absence of the MAP2K5 germline variants c.G961A and c.T1100C in a wide series of familial nonmedullary thyroid carcinoma Italian families. *International Journal of Cancer* 2019 145 600–600. (https://doi.org/10.1002/ijc.33244)

29 Srivastava A, Giangiobbe S, Skopelitou D, Miao B, Paramasivam N, Diguigiovanni C, Bonora E, Hemminki K, Forsti A & Bandapalli OR.
Whole genome sequencing prioritizes CHEK2, EWSR1, and TIAM1 as possible predisposition genes for familial non-medullary thyroid cancer. Frontiers in Endocrinology 2021 12 600682. (https://doi.org/10.3389/fendo.2021.600682)

35 He H, Li W, Liyanarachchi S, Jendrzejewski J, Srinivas M, Davuluri RV, Nagy N & de la Chapelle A. Genetic predisposition to papillary thyroid carcinoma: involvement of FOXE1, TSHR, and a novel lincRNA gene, PTCSCL2. Journal of Clinical Endocrinology and Metabolism 2015 100 E164–E172. (https://doi.org/10.1210/jc.2014-2347)

38 Da Silva AMA, Maciel RMB, Da Silva MRD, Toledo SRC, De Marques IJ, Gomes I, Pojo M, Pires C, Moura MM, Cabrera R, Santos C, He H, Li W, Liyanarachchi S, Agui K, Jendrzejewski J, Jiao H, Hoag K, Wen B, et al. Ultra-rare mutation in long-range enhancer predisposes to thyroid carcinoma with high penetrance. PLoS ONE 2013 8 e61920. (https://doi.org/10.1371/journal.pone.0061920)

35 Marques IJ, Gomes I, Pojo M, Pires C, Moura MM, Cabrera R, Santos C, van [Jcken WJ], Teixeira MR, Ramalho JS, et al. Identification of SPRY4 as a novel candidate susceptibility gene for familial nonmedullary thyroid cancer. Thyroid 2021 31 1366–1375. (https://doi.org/10.1089/thy.2020.0257)

31 Naslavsky MS, Scliar MO, Yamamoto GL, Wang JYT, Zverinova S, Da Silva AM, Maciel RMB, Da Silva MRD, Toledo SRC, De Marques IJ, Gomes I, Pojo M, Pires C, Moura MM, Cabrera R, Santos C, He H, Li W, Liyanarachchi S, Agui K, Jendrzejewski J, Jiao H, Hoag K, Wen B, et al. Ultra-rare mutation in long-range enhancer predisposes to thyroid carcinoma with high penetrance. PLoS ONE 2013 8 e61920. (https://doi.org/10.1371/journal.pone.0061920)
63 Nan H, Xu M, Zhang J, Zhang M, Kraft P, Qureshi AA, Chen C, Guo Q, Hu FB, Rimm EB, et al. Genome-wide association study identifies nidogen 1 (NID1) as a susceptibility locus to cutaneous nevi and melanoma risk. Human Molecular Genetics 2011 20 2673–2679. (https://doi.org/10.1093/hmg/ddr154)

64 Li L, Zhang Y, Li N, Feng L, Yao H, Zhang R, Li B, Li X, Han N, Gao Y, et al. Nidogen-1: a candidate biomarker for ovarian serous cancer. Japanese Journal of Clinical Oncology 2015 45 176–182. (https://doi.org/10.1093/jjco/hyu187)

65 Alečković M, Wei Y, LeRoy G, Sidoli S, Liu D, Garcia BA & Kang Y. Identification of nidogen 1 as a lung metastasis protein through secretome analysis. Genes and Development 2017 31 1439–1455. (https://doi.org/10.1101/gad.301937.117)

66 Zhou S, Zhang S, Wang L, Huang S, Yuan Y, Yang J, Wang H, Li X, Wang P, Zhou L, et al. BET protein inhibitor JQ1 downregulates chromatin accessibility and suppresses metastasis of gastric cancer via inactivating RUNX2/NID1 signaling. Oncogenesis 2020 9 33. (https://doi.org/10.1038/s41389-020-0218-z)

67 Zhu D, Xie H, Li H, Cai P, Zhu H, Xu C, Chen P, Sharan A, Xia Y & Tang W. Nidogen-1 is a common target of microRNAs MiR-192/215 in the pathogenesis of Hirschsprung’s disease. Journal of Neurochemistry 2015 134 39–46. (https://doi.org/10.1111/jnc.13118)

68 Vaes N, Schonkeren SL, Rademakers G, Holland AM, Koch A, Gijbels MJ, Keulers TG, de Wit M, Moonen L, Van der Meer JRM, et al. Loss of enteric neuronal Ndrg4 promotes colorectal cancer via increased release of Nid1 and Fbln2. EMBO Reports 2021 22 e51913. (https://doi.org/10.15252/embr.202051913)

69 Wright MT, Kouba L & Plate L. Thyroglobulin interactome profiling defines altered proteostasis topology associated with thyroid dyshormonogenesis. Molecular and Cellular Proteomics 2020 20 100008. (https://doi.org/10.1074/mcp.RA120.002168)

Received in final form 24 November 2021
Accepted 22 December 2021
Accepted Manuscript published online 23 December 2021