Molecular and Cellular Endocrinology 404 (2015) 102–112

Contents lists available at ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Novel compound heterozygous *Thyroglobulin* mutations c.745+1G>A/c.7036+2T>A associated with congenital goiter and hypothyroidism in a Vietnamese family. Identification of a new cryptic 5′ splice site in the exon 6

Cintia E. Citterio [a,b], Cecilia M. Morales [a,b], Natacha Bouhours-Nouet [c], Gloria A. Machiavelli [a,b], Elena Bueno [d], Frédérique Gatelais [c], Regis Coutant [c], Rogelio González-Sarmiento [d], Carina M. Rivolta [a,b,d], Héctor M. Targovnik [a,b,d,*]

[a] *Laboratorio de Genética y Biología Molecular, Instituto de Inmunología, Genética y Metabolismo (INIGEM, CONICET-UBA), Hospital de Clínicas "José de San Martín", C1120AAR Buenos Aires, Argentina*
[b] *Cátedra de Genética y Biología Molecular (FFyB-UBA), C1113AAD Buenos Aires, Argentina*
[c] *Unité Endocrinologie Diabétologie Pédiatrique and Centre des Maladies Rares de la Réceptivité Hormonale, CHU-Angers, 49933 Angers CEDEX 9, France*
[d] *Unidad de Medicina Molecular, Departamento de Medicina, Facultad de Medicina, Universidad de Salamanca, 37007 Salamanca, España*

ARTICLE INFO

Article history:
Received 30 August 2014
Received in revised form 17 December 2014
Accepted 21 January 2015
Available online 26 January 2015

Keywords:
Congenital hypothyroidism
Thyroglobulin gene
Mutation
Compound heterozygous mutations
Cryptic splice site

ABSTRACT

Several patients were identified with dyshormonogenesis caused by mutations in the *thyroglobulin (TG)* gene. These defects are inherited in an autosomal recessive manner and affected individuals are either homozygous or compound heterozygous for the mutations. The aim of the present study was to identify new *TG* mutations in a patient of Vietnamese origin affected by congenital hypothyroidism, goiter and low levels of serum *TG*. DNA sequencing identified the presence of compound heterozygous mutations in the *TG* gene: the maternal mutation consists of a novel c.745+1G>A (g.IVS6 + 1G>A), whereas the hypothetical paternal mutation consists of a novel c.7036+2T>A (g.IVS40 + 2T>A). The father was not available for segregation analysis. Ex-vivo splicing assays and subsequent RT-PCR analyses were performed on mRNA isolated from the eukaryotic-cells transfected with normal and mutant expression vectors. Minigene analysis of the c.745+1G>A mutant showed that the exon 6 is skipped during pre-mRNA splicing or partially included by use of a cryptic 5′ splice site located to 55 nucleotides upstream of the authentic exon 6/intron 6 junction site. The functional analysis of c.7036+2T>A mutation showed a complete skipping of exon 40. The theoretical consequences of splice site mutations, predicted with the bioinformatics tool NNSplice, Fsplice, SPL, SPLM and MaxEntScan programs were investigated and evaluated in relation with the experimental evidence. These analyses predicted that both mutant alleles would result in the abolition of the authentic splice donor sites. The c.745+1G>A mutation originates two putative truncated proteins of 200 and 1142 amino acids, whereas c.7036+2T>A mutation results in a putative truncated protein of 2277 amino acids. In conclusion, we show that the c.745+1G>A mutation promotes the activation of a new cryptic donor splice site in the exon 6 of the *TG* gene. The functional consequences of these mutations could be structural changes in the protein molecule that alter the biosynthesis of thyroid hormones.

1. Introduction

Thyroglobulin (TG) is a large homodimeric glycoprotein matrix that is required for thyroid hormone biosynthesis in thyroid cells (Targovnik, 2012). In human, it is coded by a single copy gene, 270 kb long that maps on chromosome 8q24 and contains an 8.5 kb coding sequence divided into 48 exons, with exonic sizes ranging between 63 and 1101 nucleotides separated by introns varying in size up to 64 kb (Targovnik, 2012; Targovnik et al., 2010a, 2011; van de Graaf et al., 2001).

TG mRNA is very heterogeneous due to single nucleotide polymorphisms, polyadenylation cleavage site variants and alternatively spliced transcripts (Targovnik, 2012; van de Graaf et al., 2001). The excision of introns and the joining of exons, which is defined as constitutive and alternative splicing, is a complex process in the *TG*

* Corresponding author. Laboratorio de Genética y Biología Molecular, Instituto de Inmunología, Genética y Metabolismo (INIGEM, CONICET-UBA), Hospital de Clínicas "José de San Martín", Av. Córdoba 2351, Cuarto Piso, Sala 5, C1120AAR Buenos Aires, Argentina. Tel: +54 11 5950 8805; fax: +54 11 4964 8296.
E-mail address: htargovnik@conicet.gov.ar; htargovn@ffyb.uba.ar (H.M. Targovnik).

http://dx.doi.org/10.1016/j.mce.2015.01.032
0303-7207/

gene due to their size. Importantly, alternative splicing is a major contributing factor in eukaryotic cells, which allows a relatively small number of genes to produce several isoforms that lead to different proteins (Gamazon and Stranger, 2014). Unfortunately, many aspects of the alternative splicing mechanism of *TG* in normal and pathological conditions have not been understood. At present, the alternative splicing map of human *TG* mRNA consists of a series of 11 transcripts where complete exons or a group of consecutive complete exons were skipped (Bertaux et al., 1991, 1995; Hishinuma et al., 1999; Mason et al., 1995; Targovnik et al., 1992; van de Graaf et al., 1999). The regions absent in each of these mRNA molecules correspond to exons 3, 4 and 6, 17, 17 and 18, 17–19, 17–20, 22, 26 and 46 (Bertaux et al., 1991, 1995; Hishinuma et al., 1999; Mason et al., 1995; Targovnik et al., 1992; van de Graaf et al., 1999). Two deletions of nucleotides between positions 3430–3736 and 7301–7561 with partial exclusion of exons were also identified (van de Graaf et al., 1999). Previous studies have reported that the human p.R1511X mutation in exon 22 (Caputo et al., 2007a, 2007b; Gutnisky et al., 2004; Targovnik et al., 1993, 2010b) together with the p.R687X mutation located in the exon 9 of the Afrikander cattle (Ricketts et al., 1987) illustrates the phenomenon of nonsense-mediated altered splicing, in which an exon harboring a premature stop codon is removed from the mature transcript by exon skipping. These alternative splicings are also present in mRNA from normal thyroid tissue, but they represent a very minor fraction of the total *TG* transcripts.

The exon/intron boundaries, known as the 5′ and 3′ splice sites (ss) or donor and acceptor ss respectively, are defined by consensus motifs that are critical for the splicing machinery (Roca et al., 2003). The most common consequence of splicing mutations is skipping of one or more exons, followed by activation of cryptic donor and acceptor sequences, whereas intron retention is very rare. (Buratti et al., 2007; Roca et al., 2003).

Several patients were identified with dyshormonogenesis caused by mutations in the *TG* gene (Abdul-Hassan et al., 2013; Agretti et al., 2013; Alzahrani et al., 2006; Cangul et al., 2014; Caputo et al., 2007a, 2007b; Caron et al., 2003; Citterio et al., 2011, 2013a, 2013b; Corral et al., 1993; González-Sarmiento et al., 2001; Gutnisky et al., 2004; Hermanns et al., 2013; Hishinuma et al., 1999, 2005, 2006; Ieiri et al., 1991; Kahara et al., 2012; Kanou et al., 2007; Kim et al., 2008; Kitanaka et al., 2006; Liu et al., 2012; Machiavelli et al., 2010; Moya et al., 2011; Narumi et al., 2011; Niu et al., 2009; Pardo et al., 2008, 2009; Pérez-Centeno et al., 1996; Peteiro-Gonzalez et al., 2010; Raef et al., 2010; Rivolta et al., 2005; Targovnik et al., 1993, 1995, 2001, 2010b, 2012; van de Graaf et al., 1999). These defects are inherited in an autosomal recessive manner and affected individuals are either homozygous or compound heterozygous for the mutations. Fourteen ss mutations have been identified in the human *TG* causing exon skipping events, involving the +1/+2/+3/+4 positions in donor ss relative to the exon-intron junction (c.638+1G>A (g.IVS5+1G>A; Alzahrani et al., 2006; Abdul-Hassan et al., 2013), c.4932+1G>C (g.IVS24+1G>C; Hishinuma et al., 2006), c.5686+1G>A (g.IVS30+1G>A; Hishinuma et al., 2006), c.5686+1G>T (g.IVS30+1G>T; Targovnik et al., 1995, 2001; Pardo et al., 2008, 2009), c.5686+1G>C (g.IVS30+1G>C; Hermanns et al., 2013), c.6262+1delG (g.IVS35 + 1delG; Peteiro-Gonzalez et al., 2010), c.274+2T>G (g.IVS3+2T>G; Niu et al., 2009), c.7862+2T>A (g.IVS45+2T>A; Hishinuma et al., 2006), c.4159 + 3_+4delAT (g.IVS19+3_+4delAT; Targovnik et al., 2012)) or −3/−2/−1 position in acceptor ss relative to the intron–exon junction (c.177−3C>G (g.IVS3−3C>G; Ieiri et al., 1991), c.6398−2A>G (g.IVS37−2A>G; Narumi et al., 2011), c.2177−1G>A (g.IVS10−1G>A; Hishinuma et al., 2006; Abdul-Hassan et al., 2013), c.6056−1G>C (g.IVS34−1G>C; Gutnisky et al., 2004; Abdul-Hassan et al., 2013), c.7863−1G>A (g.IVS46−1G>A; Pardo et al., 2009)). Our lab has recently identified the first cryptic ss in the *TG* gene (Targovnik et al., 2012). We have demonstrated that the c.4159 + 3_+4delAT mutation promotes the activation of cryptic donor ss in the exon 19 (Targovnik et al., 2012).

In the present study we report a patient with CH and low levels of serum *TG*. Screening by direct sequencing analysis of the *TG* gene revealed two novel ss mutations, conforming compound heterozygous mutations, c.745+1G>A/c.7036+2T>A (g.IVS6+1G>A/g.IVS40+2T>A). Remarkable, the donor ss mutation in intron 6 results in the skipping of the exon 6 and in the activation of an exonic cryptic 5′ ss with the consequent deletion of 55 bp at the end of this exon. To our knowledge this is the first time that a mutation in the *TG* gene is reported in a congenital hypothyroid Vietnamese patient.

2. Materials and methods

2.1. Patient

We report a boy born in Vietnam in 1998, from non consanguineous parents and affected with congenital goitrous hypothyroidism. He had no siblings. Neonatal history could not be obtained. He has been living in France since 2003 and was referred to the endocrinology centre for the management of his goiter and his hypothyroidism. He began treatment with thyroid hormone replacement.

At the age of 8 years and 3 months he was clinically re-evaluated under treatment with 112.5 μg of levothyroxine per day, his weight was 27.3 kg (+0.7 DS) and his height was 131.5 cm (+1 DS). Bone age was 7 years and 10 months. Clinically, we found a goiter with normal consistence and without nodules. Biochemical thyroid parameters under 112.5 μg of levothyroxine indicated elevated TSH of 68.17 mUI/l (reference values: 0.35–3.5) with low Free T_4 5.7 pmol/l (reference values: 7.5–21.1) and normal Free T_3 5.5 pmol/l (reference values: 3.8–6). Negative anti *TPO* (<10 U/ml), anti *TG* (<2.2 UI/ml) and anti *TSH* receptor (<0.3 UI/l) antibodies. The serum *TG* concentration was low at <0.3 ng/ml (reference values: < 15 ng/ml), suggesting that hypothyroidism could be related to defective *TG* synthesis. Thyroid ultrasound showed an enlargement of the gland of 23.2 ml with homogeneous pattern. Thyroid volume was calculated by multiplication of length, breadth and depth and a corrective factor of 0.52 for each lobe. Reference range for 6–11 years: 2.7 ± 0.8 ml (Citterio et al., 2013a). Scintigraphy showed a normally located thyroid gland and his potassium perchlorate discharge test was negative (4%). Levothyroxine dose was increased to 125 μg per day with the consequent normalization of thyroid parameters: TSH 0.25 mUI/l, Free T4 18.9 pmol/l, Free T3 4.9 pmol/l. The goiter remained without changes on successive ultrasound examinations during follow-up.

His mother (I-2) had normal biochemical thyroid parameters, TSH 1.7 mUI/l (reference values: 0.27–4.2), Free T_4 16.4 pmol/l (reference values: 12–22), Free T_3 5.1 pmol/l (reference values: 3.1–6.8), negative anti *TG* (<0.9 UI/ml), anti *TPO* (<10 U/ml) and anti *TSH* receptor (<0.3 UI/l) antibodies. We have no detail about his biological father who is still living in Vietnam.

Written informed consent was obtained from the mother and the research project was approved by the Institutional Review Board.

2.2. DNA sequencing

Genomic DNA was isolated from peripheral blood leucocytes by using standard methods and stored at −20 C until analyzed.

The 180 bp of the promotor region and all 48 exons of the *TG* gene, including splicing signals and the flanking intronic regions were amplified by PCR and sequenced using sense and antisense specific primers or M13 universal primers reported previously (Gutnisky et al., 2004), with the Big Dyedeoxyterminator Cycle Sequencing Kit (Applied Biosystems, Weiterstadt, Germany). The samples were analyzed on the ABI Prism 3100 DNA sequencer (Applied Biosystems).

2.3. Cloning of wild-type and mutated exon 17 PCR fragments

The amplified fragment corresponding to exon 17 from the mother of the index patient II-1 was T-A cloned into pGEM-T Easy vector (Promega, Madison, WI, USA). DNA sequencing was performed as described above from wild type and mutant allele clones using T7 vector primer.

2.4. Validation of c.3665C>T [p.S1203L] and c.3808C>T [p.R1251C] mutations by Single-Strand Conformation Polymorphism (SSCP) analysis

We validated the c.3665C>T [p.S1203L] and c.3808C>T [p.R1251C] mutations studying healthy unrelated individuals by SSCP using intronic primers previously reported to amplify exon 17 (Gutnisky et al., 2004). The gel matrix for SSCP analysis contained 10% polyacrylamide (29:1) without glycerol. Samples were electrophoresed for 24 h at a constant temperature (4 °C). DNA was visualized by silver-staining.

2.5. Construction and expression of the minigenes

To study the effect of the c.745+1G>A and c.7036+2T>A mutations, we constructed wild-type and mutated hybrid minigenes using the vector pSPL3 (Life Technologies Inc., Gaithersburg, MD, USA). The two genomic DNA regions from index patient II-1 were amplified by long PCR technique using elongase (Invitrogen, Carlsbad, CA, USA), one region containing exon 6 (107-bp) and their intronic flanking sequences (502-bp upstream from the 5′ exon end and 366-bp downstream from the 3′ exon end), and the other region containing exon 40 (160-bp) and their neighboring intronic sequences (1024-bp upstream from the 5′ exon end and 1044-bp downstream from the 3′ exon end). The oligonucleotides used to amplify the two genomic DNA regions are as follows: Intron 5/Exon 6/Intron 6 region: forward primer (SPL3-I5-F), 5′-ATAAGAAT GCGGCCGCTAAGGATTTCCTGTGGTGGC-3′ and reverse primer (SPL3-I6-R), 5′-CAGGATCCCAGTTTCCAAAGTCAGAGCTGC-3′; Intron 39/Exon 40/Intron 40 region: forward primer (SPL3-I39-F), 5′-ATAAGAATGCGGCCGCCTTGGATGGAGCACGAGTTA-3′ and reverse primer (SPL3-I40-R), 5′-CAGGATCCTGTGGTCATCCAGGTTTCTG-3′. The forward primers contained the *NotI* site and the reverse primers contained the *BamHI* site (underlined). The amplified products of 999-bp (975 of which were Intron 5/Exon 6/Intron 6 *TG* sequences) and of 2252-bp (2228 of which were Intron 39/Exon 40/Intron 40 *TG* sequences) were sequentially digested with *BamHI* and *NotI*. The insert was directionally cloned into the *NotI* and *BamHI* sites of the pSPL3 vector. Because index patient II-1 is heterozygous for the both mutations, the wild-type and mutated alleles were thus cloned. The recombinant plasmids were amplified in DH5α-competent cells and purified by use of the Wizard Plus SV Minipreps DNA Purification System (Promega). The correct sequence was confirmed by sequencing with the respective intronic (TG6F, TG6R, TG39F, TG39R (Gutnisky et al., 2004)) and cloning (SPL3-I5-F, SPL3-I6-R, SPL3-I39-F, SPL3-I40-R) primers.

Cos1, CV1 and HeLa cells were grown in 3.8-cm dishes in DMEM supplemented with 10% fetal bovine serum and 100 U/ml penicillin-streptomycin (Invitrogen) in a 5% CO_2 atmosphere at 37 °C. When cells reached approximately 80% confluence, they were transfected with 500 ng of plasmid DNA (wild-type, mutant, and control pSPL3)/3.8-cm dish with the Lipofectamine 2000 (Invitrogen). Twenty four hours later, cells were harvested and total RNA was extracted with SV Total RNA Isolation System (Promega).

The RT-PCR was performed as described previously (Targovnik et al., 1993) using vector-specific primers: forward primer (pSPL3F), 5′-TCTGAGTCACCTGGACAACC-′3 and reverse primer (pSPL3R), 5′-ATCTCAGTGGTATTTGTGAGC-3′. Samples were heated to 94 °C for 5 min, followed by 39 cycles of DNA denaturation (94 °C for 30 s), annealing (55 °C for 30 s), and polymerization (72 °C for 1 min). After the last cycle, the samples were incubated for an additional 10 min at 72 °C. The RT-PCR products were purified from the agarose gel by using the Illustra GFX PCR DNA and Gel Band Purification Gel (GE Healthcare, Buckinghamshire, UK), cloned into pGEM-easy vector (Promega) and then sequenced with the T7 primer.

2.6. Computer prediction analysis

Searching for potential 5′ ss sequences in the *TG* gene spanning from exon 6 to intron 6 and exon 40 to intron 40 was accomplished using the NNSplice (http://www.fruitfly.org/seq_tools/splice.html), Fsplice (http://linux1.softberry.com/berry.phtml?topic=fsplice&group=programs&subgroup=gfind), SPL (http://linux1.softberry.com/berry.phtml?topic=spl&group=programs&subgroup=gfind) and SPLM (http://linux1.softberry.com/berry.phtml?topic=splm&group=programs&subgroup=gfind) prediction tools. Scores of the 9-nt sequences, corresponding either to authentic, mutated and cryptic 5′ ss, were calculated by means of the MaxEntScan program (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) by selecting for the analyses the MAXENT, the MDD, the MM, or the WMM. Finally, the analysis of exon sequences that allow the identification of putative exonic splicing enhancer (ESE) responsive to the human Ser/Arg-rich proteins (SR proteins) (Blencowe, 2000) was performed using the ESEfinder program (http://rulai.cshl.edu/tools/ESE/).

2.7. Nucleotide, amino acid and splicing mutations nomenclatures

The nucleotide position in human *TG* mRNA is designated according to reference sequences (GenBank accession number NM_003235). The A of the ATG of the initiator methionine codon is denoted nucleotide +1. The amino acid positions are numbered after subtracting the 19-amino-acid signal peptide. Splicing mutations are annotated by using cDNA sequences and old nomenclature (g.IVS) is included in parenthesis.

3. Results

3.1. DNA sequence analysis of TG

All 48 exons of the *TG* gene from index patient II-1 and his mother were screened as well as 180 bp of the *TG* promoter and all the flanking intronic sequences by direct DNA sequencing. A total of 15,000 bases were analyzed. The GT-AG splicing consensus sequences were rigorously respected in all introns of the gene under study in the propositus, except for the donor ss of introns 6 and 40. We identified a novel heterozygous guanine-to-adenine transition at the invariant +1 position of the donor of intron 6 c.745+1G>A (g.IVS6+1G>A) and a novel heterozygous thymine-to-adenine transversion at the invariant +2 position of the donor of intron 40 c.7036+2T>A (g.IVS40+2T>A) (Fig. 1). This finding indicates that the II-1 is a compound heterozygous for c.745+1G>A/c.7036+2T>A mutations in the *TG* gene. The healthy mother, who has normal thyroid function, was heterozygous for the c.745+1G>A mutation and did not carry the c.7036+2T>A mutation (Fig. 1). Unfortunately, the father was not available for segregation analysis. However, it is conceivable to hypothesize that the propositus has inherited one copy of the c.7036+2T>A mutation from his father, without excluding the possibility that it was a de novo mutational event.

In addition, direct sequencing of PCR product of exon 17 from the mother revealed two heterozygous variants of sequence, c.3665C>T [p.S1203L] and c.3808C>T [p.R1251C], whereas index patient II-1 was found to be heterozygous carrier of c.3665C>T and homozygous carrier of c.3808C (Fig. 1). An important observation

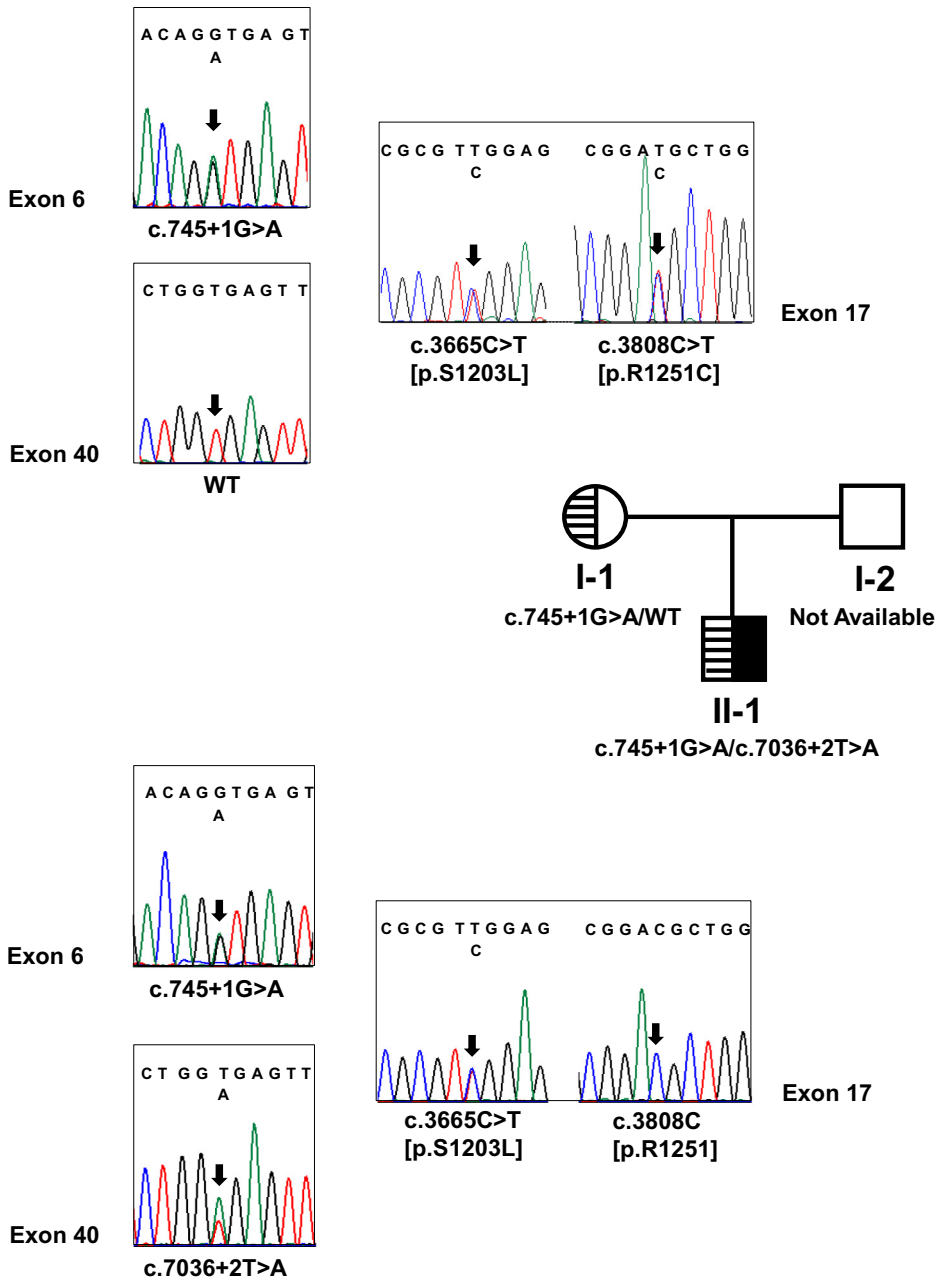

Fig. 1. Mutations in the thyroglobulin gene in index patient II-1 and his mother. Partial sequencing chromatograms of genomic DNA are shown. The pedigree shows the pattern of inheritance of the mutant thyroglobulin alleles. Squares represent males and circles females. Filled symbols denote affected individuals and half-filled symbols, unaffected heterozygous individuals. The hatched symbols indicate the c.745+1G>A mutated allele and the solid symbols the c.7036+2T>A mutated allele. Sense strand is shown. Arrows denote the position of identified mutations and variation of sequences in exon 17 (c.3665C>T and c.3808C>T), single chromatogram peaks indicate homozygosity and two overlapping peaks at the same locus, heterozygosity. Note that the index patient II-1 have inherited one copy of the c.745+1G>A mutation from his mother (I-1) and hypothetically one copy of the c.7036+2T>A mutation from his father (I-2). The DNA of the father was not available.

is that pGemT-easy vector cloning of mother's exon 17 and subsequent sequence shows that two identified variations, c.3665T and c.3808T, are located in different chromosomes. This implies that the c.745+1G>A mutation is associated with c.3665T and c.3008C alleles, and c.7036+2T>A mutation is associated with c.3665C and c.3008C alleles.

3.2. Validation of the c.3665C>T [p.S1203L] and c.3808C>T [p.R1251C] variants of sequences by SSCP analysis

The c.3665C>T and c.3808C>T variants of sequences were not found in DNA samples from 50 healthy unrelated control subjects by SSCP analysis, indicating that this variations occurs in less than 1% of the general population.

3.3. Minigene analysis

To assess the impact of the ss mutations (c.745+1G>A/c.7036+2T>A), minigene constructs were generated and tested in transiently transfected cultured cells by RT-PCR amplification. Wild-type and mutated RT-PCR fragments were cloned in pGemT-easy vector and then sequenced with the T7 primer.

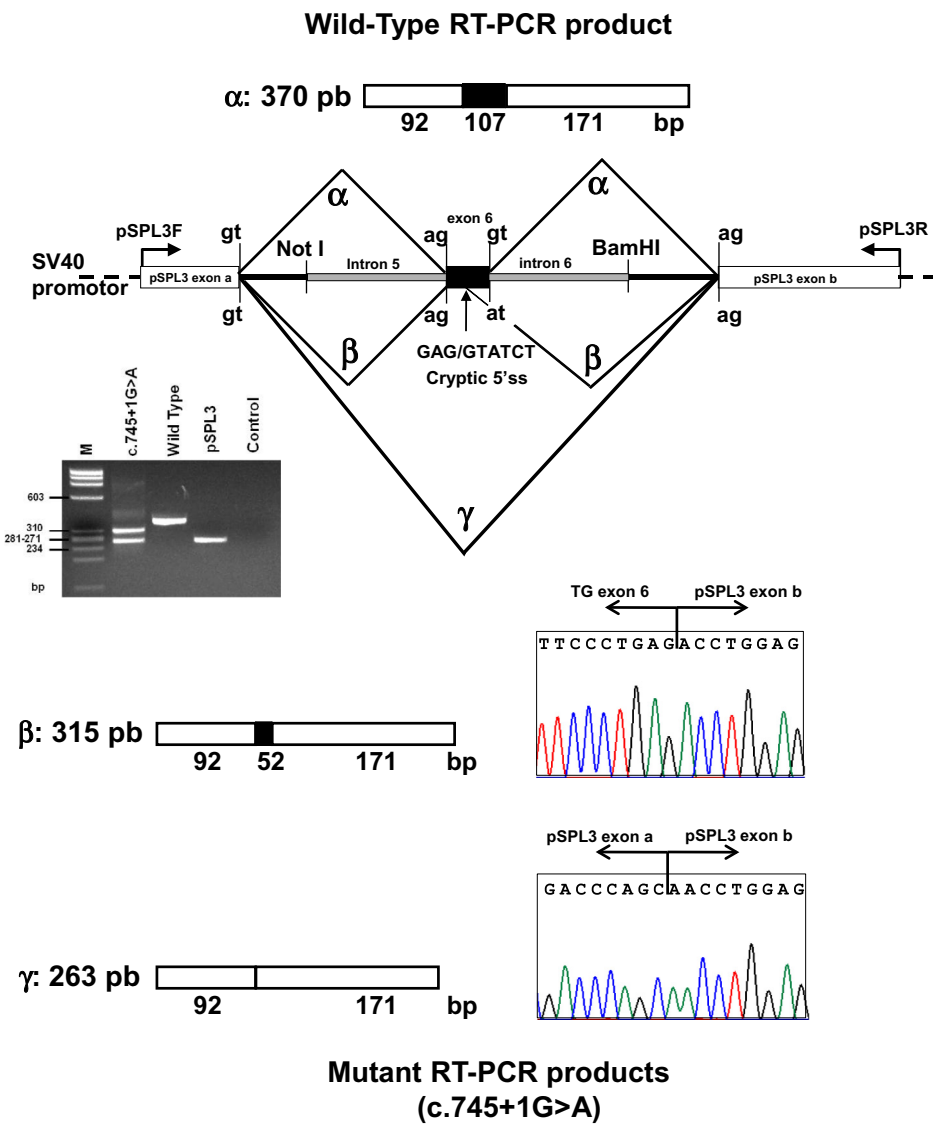

Fig. 2. In vitro expression of the wild-type and c.745+1G>A mutant minigenes. Schematic representation of the genomic organization of the wild-type and mutant minigenes and their RT-PCR products. The 999-bp PCR amplified fragments from index patient II-1 were directionally cloned into the NotI and BamHI sites of the exon trapping pSPL3 vector, which was expressed in eukaryotic cells, grown in 3.8-cm dishes. Vector and genomic DNA splice donor (GT) and acceptor (AG) sites are shown. cDNA was synthesized from transcribed mRNA and amplified with pSPL3F and pSPL3R primers complementary to flanking vector sequences. The processing of the wild-type transfected sequences (α splicing event) produces a RT-PCR product of 370 bp. In contrast, the c.745+1G>A mutation results in the activation of cryptic 5′ splice site in the exon 6 (β splicing event) and the skipping of exon 6 (γ splicing event), producing a 315 bp (55 bp deletion at the end of exon 6) and a 263 bp fragments, respectively. Gel electrophoresis of the RT-PCR amplification products (c.745 + 1G>A mutant, wild-type, empty pSPL3 vector and negative control without RNA) and partial sequencing chromatograms corresponding to β and γ splicing events are shown. The size marker is a ΦX174 DNA-Hae III ladder (M).

3.3.1. c.745+1G>A construct

Minigene assays revealed that the c.745+1G>A mutation created two different transcripts: one of them corresponds apparently to the partially included exon 6 (β splicing event) and the smallest transcript is compatible in size with the skipping of exon 6 (γ splicing event) (Fig. 2), whereas the minigene containing the wild-type allele only showed one band, that would correspond to the correct transcript (α splicing event) (Fig. 2). Sequence analysis confirmed the exclusion of exon 6 or the deletion of 55 nucleotides at the end of exon 6 in the mutated transcripts and that the wild-type transcript contained the correct exon 6 (Fig. 2). The internal deletion of exon 6 was anticipated as result from the described 5′ splice site mutation, which activates a cryptic site resembling the 5′ ss consensus sequence. The nucleotide composition of cryptic 5′ ss activated is GAG/GTATCT (backslashes indicate the cryptic exon–intron junction site), 55 nucleotides upstream of the authentic exon–intron junction site (Fig. 2).

3.3.2. c.7036+2T>A construct

The in vitro experiments showed that the processing of the c.7036+2T>A minigene generated a fragment with a size that would correspond to the skipping of involved exon 40 (β splicing event) (Fig. 3). The wild type minigene showed one fragment, corresponding to the expected size with the correct processing of exon 40 (α splicing event) (Fig. 3). DNA sequencing confirmed the exclusion of exon 40 in the mutated transcripts and that the wild-type transcript contained the correct exon 40 (Fig. 3).

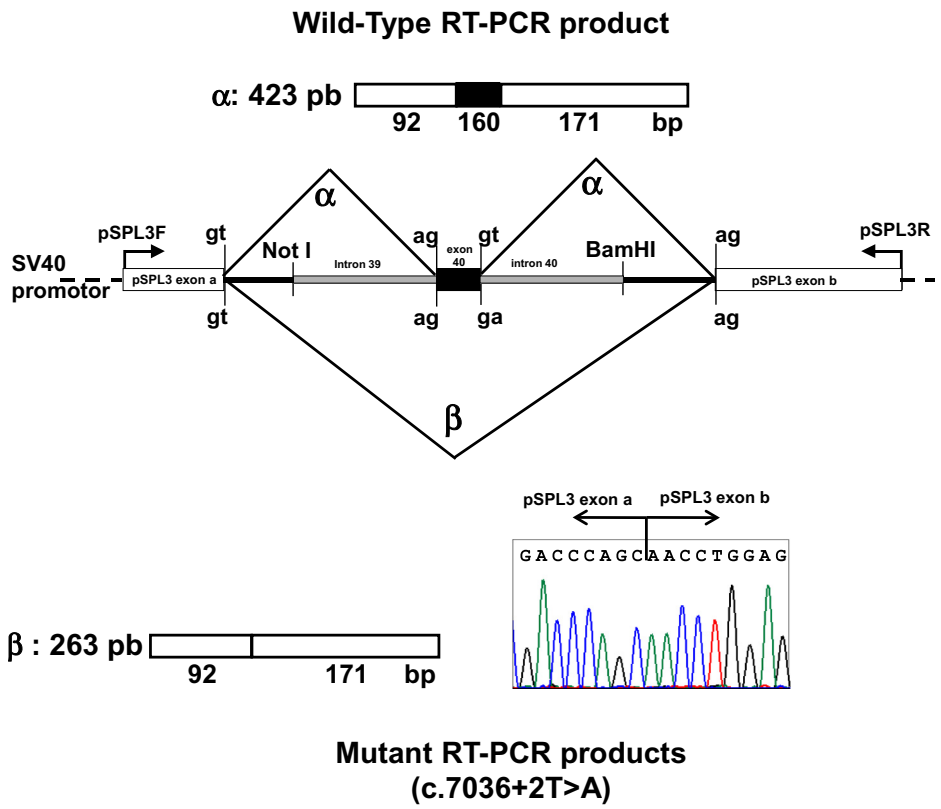

Fig. 3. In vitro expression of the wild-type and c.7036+2T>A mutant minigenes. Schematic representation of the genomic organization of the wild-type and mutant minigenes and their RT-PCR products. The 2252-bp PCR amplified fragments from index patient II-1 were directionally cloned into the NotI and BamHI sites of the exon trapping pSPL3 vector, which was expressed in eukaryotic cells, grown in 3.8-cm dishes. Vector and genomic DNA splice donor (GT) and acceptor (AG) sites are shown. cDNA was synthesized from transcribed mRNA and amplified with pSPL3F and pSPL3R primers complementary to flanking vector sequences. The processing of the wild-type transfected sequences (α splicing event) produces a RT-PCR product of 423 bp. In contrast, the c.7036+2T>A mutation results in the skipping of exon 40 (β splicing event), producing a 263 bp fragment. Partial sequencing chromatogram corresponding to β splicing event is shown.

3.4. 5′Splice site prediction analysis

In order to evaluate the in silico relevance of the cryptic 5′ ss sequence, the potential 5′ ss located in the exon 6, upstream of the wild type 5′ ss was predicted using the NNSplice, FSplice, SPL and SPLM tools. As shown in Fig. 4a, the exonic cryptic 5′ ss activated, predicted by minigene analysis, was recognized by all programs, except for SPL. The strength of cryptic and physiologic sites were also compared by four other methods, the maximum entropy model (MAXENT), the maximum dependence decomposition model (MDD), the first-order Markov model (MM) and Weight Matrix Model (WMM), that require the prior knowledge of the input sequence to be tested. MAXENT, MDD, MM and WMM scores were consistently lower for the cryptic 5′ ss with respect to wild type 5′ ss (Fig. 4a). As expected, NNSplice, FSplice, SPL and SPLM combined prediction program analysis did not identify the mutated 5′ ss as a donor site of splicing (Fig. 4a). Using the ESEfinder program we observed that the cryptic site has ESE sequences recognizable by SF2/ASF, SC35, SRp40 and SRp55 proteins, whereas wild type 5′ ss was recognized by SF2/ASF and SRp40, and mutated 5′ ss only by SF2/ASF (Fig. 4b).

In-silico prediction of the effect of the c.7036+2T>A mutation on splicing showed the expected inactivation of intron 40 donor site (Fig. 5). The location of such substitutions usually has an adverse effect on the recognition of ss by the cellular machinery. NNSplice, FSplice, SPL and SPLM combined prediction program analysis did not identify the mutated 5′ ss as a donor site of splicing (Fig. 5a), whereas MAXENT, MDD, MM and WMM scores were much lower for the mutated 5′ ss with respect to wild type 5′ ss (Fig. 5a). Interestingly, according to ESEfinder, this mutation does not diminish the score or eliminates the SR proteins that bind to ESE sequences located around the exon/intron 40 junction (Fig. 5b).

3.5. TG proteins prediction analysis

According to the results obtained by expression of the minigenes, the aberrant splicing by c.745+1G>A mutation induces a reading frame change by junction of the exon 5 or partially deleted exon 6 with the exon 7, generating two truncated proteins (Fig. 6). Therefore, the skipping of exon 6 results in a frameshift at amino acid 195 and aberrant splicing cause the activation of a cryptic 5′ ss that produces a frameshift at amino acid 212 (Fig. 6). Both reading frame changes give rise to their respective putative premature stop codons, 6 codons downstream, within exon 7, in skipping exon 6 or 152 codons downstream, within exon 9, in partially deleted exon 6 (Fig. 6). Consequently, the finally encoded protein from the total deleted exon 6 is predicted to contain 200 residues, including 194 residues of normal *TG* sequences, and from the partial deleted exon 6 is predicted to contain 1142 residues, including 211 residues of normal *TG* sequence (Fig. 6).

The absence of exon 40 by c.7036+2T>A mutation results in a frameshift at amino acid 2274 with a putative premature stop codon 4 codons downstream, located in the same exon 41 (Fig. 7). Consequently, the finally encoded protein is predicted to contain 2277 residues, including 2273 residues of normal *TG* sequence (Fig. 7).

a

	Cryptic 5'ss GAG/GTATCT	Wild-type 5'ss CAG/GTGAGT	Mutated 5'ss CAG/ATGAGT
NNsplice	0.37	0.99	0.00
FSplice	6.80	13.38	0.00
SPL	0.00	0.92	0.00
SPLM	8	94	0
MAXENT	7.27	10.67	2.49
MDD	9.48	14.88	6.70
MM	5.40	11.87	3.69
WMM	5.09	11.69	3.51

b

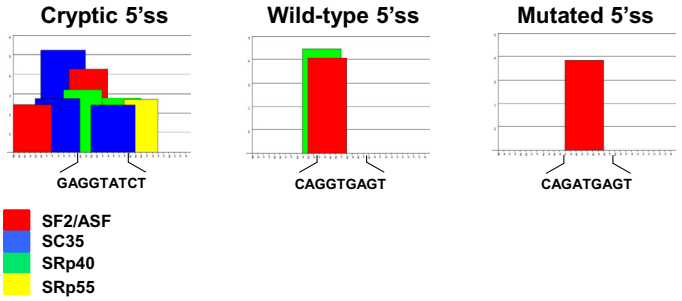

Fig. 4. c.745+1G>A mutation, *in silico* analysis of the cryptic, wild-type and mutated 5′ splice sites. (a) Individual scores for each 5′ splice sites obtained by a panel of donor site prediction programs (NNSplice, FSplice, SPL, SPLM MAXENT, MDD, MM and WMM). (b) Potential SR binding sites (SF2/ASF, SC35, SRp40 and SRp55) identified by the ESEfinder 3.0 program. The height of each bar represents the motif scores, whereas its width indicates the length of the binding site motifs for SR proteins and its position along the sequence. Note that the cryptic 5′ splice site is potentially recognizable by SR proteins. / denotes the exon/intron boundary.

a

	Wild-type 5'ss CTG/GTGAGT	Mutated 5'ss CTG/GAGAGT
NNsplice	0.87	0.00
FSplice	12.54	0.00
SPL	0.86	0.00
SPLM	86	0
MAXENT	10.10	1.91
MDD	14.48	6.29
MM	9.36	1.18
WMM	9.66	1.47

b

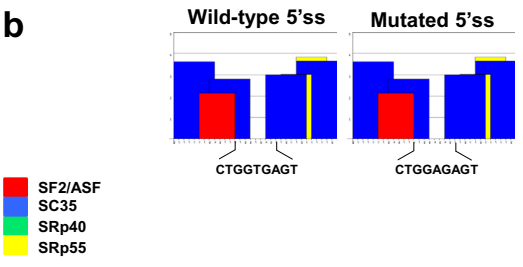

Fig. 5. c.7036+2T>A mutation, *in silico* analysis of the wild-type and mutated 5′ splice sites. (a) Individual scores for each 5′ splice sites obtained by a panel of donor site prediction programs (NNSplice, FSplice, SPL, SPLM MAXENT, MDD, MM and WMM). (b) Potential SR binding sites (SF2/ASF, SC35, SRp40 and SRp55) identified by the ESEfinder 3.0 program. The height of each bar represents the motif scores, whereas its width indicates the length of the binding site motifs for SR proteins and its position along the sequence. / denotes the exon/intron boundary.

4. Discussion

In this study, we used an ex-vivo splicing assay based on a minigene system and bioinformatics tools to examine the effects at the mRNA level of two new ss mutations, c.745+1G>A and c.7036+2T>A, in the *TG* gene. Defects in pre-mRNA splicing have been shown as a common disease-causing mechanism in a considerable number of pathophysiological entities, either by altering positions of ss sequences or by affecting intronic or exonic splicing regulatory sequences such as ESE elements (Bonnet et al., 2008; Tournier et al., 2008). In particular, a missplicing of *TG* pre-mRNA due to a mutation in consensus splice donor or acceptor site is known to induce a congenital goiter and hypothyroidism in humans (Abdul-Hassan et al., 2013; Alzahrani et al., 2006; Gutnisky et al., 2004; Hermanns et al., 2013; Hishinuma et al., 2006; Ieiri et al., 1991; Narumi et al., 2011; Niu et al., 2009; Pardo et al., 2008, 2009; Peteiro-Gonzalez et al., 2010; Targovnik et al., 1995, 2001, 2012) and rodents (Sato et al., 2014). The usefulness of splicing reporter minigene assays has been shown to be a good approach to determine the effect of the variants on the splicing process, (Bonnet et al., 2008; Tournier et al., 2008), when genes present a restricted expression profile or is difficult to obtain RNA from patients' tissues. A high level of concordance between data obtained with these assays and data from patient's RNA has been shown (Bonnet et al., 2008). Because the thyroid tissue and blood cells from II-1 were unavailable we used hybrid minigene constructs to evaluate both ss mutations. In the present study, minigenes were constructed using the pSPL3 vector, which has a minimal gene organization: the SV40 promoter followed by an exon–intron–exon structure with a multiple cloning site located inside the intron. These constructions were expressed in living cells where the splicing machinery remains intact. We found that besides exon 6 skipping, the c.745+1G>A mutation causes the activation of a cryptic 5′ ss localized in the exon 6. Cryptic ss by definition are splice sites that are not detectably used in wild-type pre-mRNA, but are only selected as a result of a mutation elsewhere in the gene, most often at the authentic ss (Roca et al., 2003). The total absence or the partial inclusion of exon 6 changes the open reading frame of the mutant transcript and generates a downstream premature translation stop codon within exon 7 or exon 9, respectively (Fig. 6). These results are identical to those previously observed by us in experimental assays showing the activation of the cryptic ss in exon 19 by c.4159+3_+4delAT mutation (Targovnik et al., 2012). Hybrid minigene analysis of the c.4159+3_+4delAT mutant showed that the exon 19 is skipped during pre-mRNA splicing or partially included by use of cryptic 5′ ss located to 100 nucleotides upstream of the wild type exon–intron junction. Consequently, the c.4159+3_+4delAT mutation originates two putative truncated proteins of 1330 and 1349 amino acids (Targovnik et al., 2012).

On the other hand, *in vitro* transcription results showed that the second ss mutation identified in the proband, c.7036+2T>A, also altered the donor site of exon 40, leading to the total skipping of this exon. The skipping of exon 40 generates a translation frameshift which introduces a premature stop codon within exon 41 (Fig. 7).

The human 5′ consensus splice donor sequence is AG/GTRAGT (R indicates purine and backslashes indicate the exon–intron junctions site) (Buratti et al., 2007; Roca et al., 2003; Shapiro and Senapathy, 1987). For both ss variants, computer-assisted ss prediction analysis recognized a significant difference between the wild-type and the mutant sequences, suggesting that they might impact *TG* splicing. Several studies have suggested that cryptic 5′ ss scores are generally lower than those calculated for the authentic ss they replace (Roca et al., 2003). The local context of those cryptic sites allows the splicing machinery to select the natural sites (Roca et al., 2003). In this context, the present study provided new evidence in

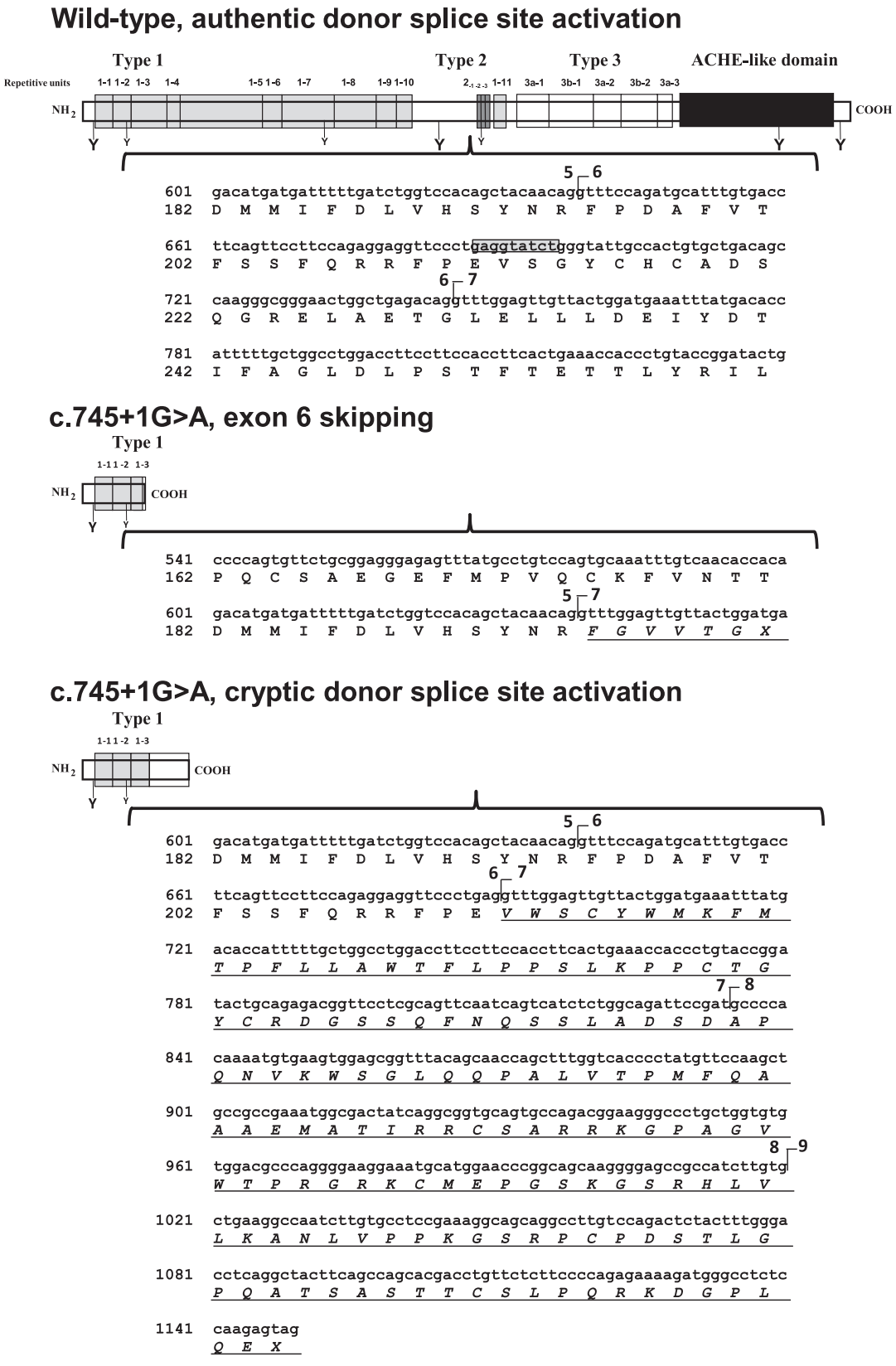

Fig. 6. Structural organization of the wild-type and putative c.745+1G>A mutant thyroglobulin proteins. The repetitive units of type 1, 2 and 3 and the acetylcholinesterase-homology domain (ACHE-like domain) are represented by boxes. The N-terminal limit of repeat type 1–5 is ambiguous. Tyrosine residues involved as acceptor (Y) and donor (ʏ) sites in thyroid hormone synthesis are shown. The partial nucleotide and the deduced amino acid sequences from wild-type and putative c.745+1G>A mutants are reported below of the respective protein diagrams. The nucleotide sequence is given in the upper line, and the amino acid translation (represented by single-letter code) is given below their respective codons. The nucleotide position in human *TG* mRNA is designated according to reference sequences (GenBank Accession No.: NM_003235). The A of the ATG of the initiator methionine codon is denoted nucleotide +1. The amino acid positions are numbered after subtracting the 19-amino-acid signal peptide. The cryptic 5′ splice site is indicated by gray shaded square and the resulting reading frame changes by cryptic 5′ splice site activation and exon 6 skipping are underlined. ⌜ indicates exon/exon boundaries and exon numbering is indicated.

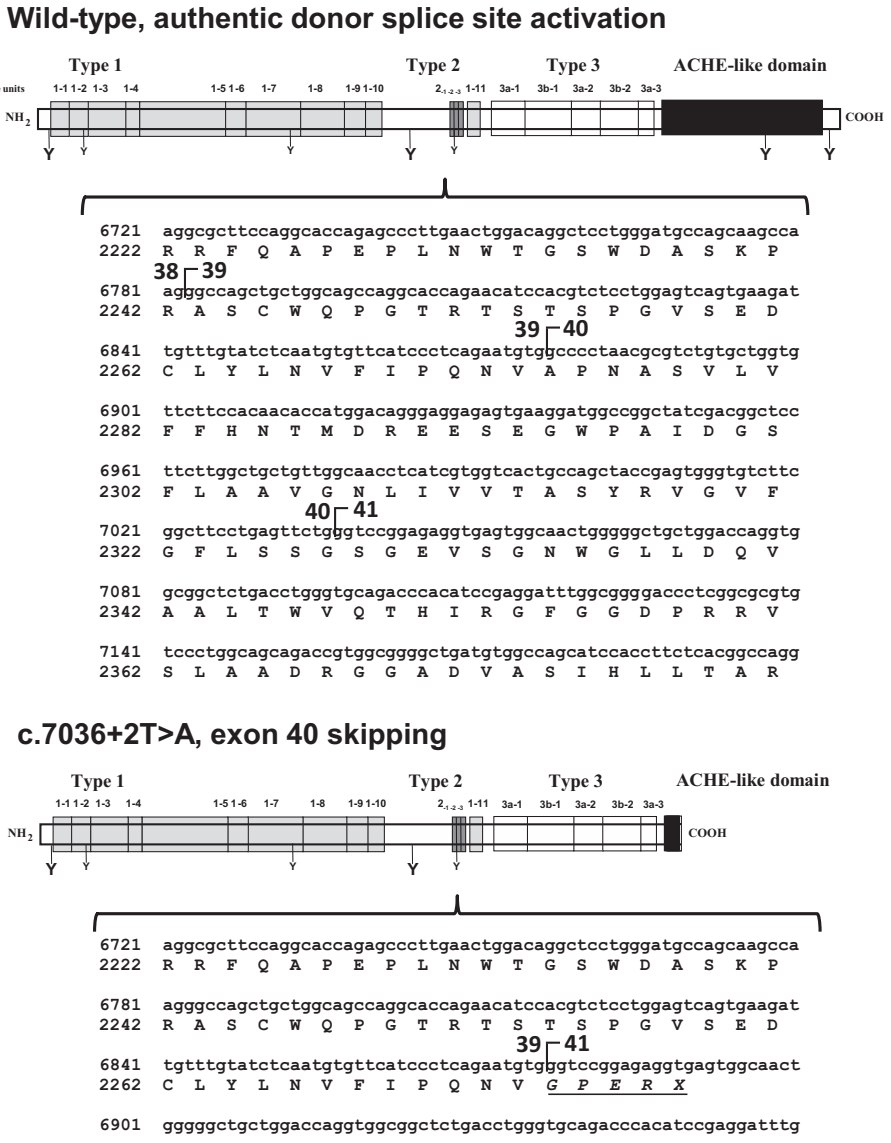

Fig. 7. Structural organization of the wild-type and putative c.7036+2T>A mutant thyroglobulin proteins. The repetitive units of type 1, 2 and 3 and the acetylcholinesterase-homology domain (ACHE-like domain) are represented by boxes. The N-terminal limit of repeat type 1–5 is ambiguous. Tyrosine residues involved as acceptor (Y) and donor (Y) sites in thyroid hormone synthesis are shown. The partial nucleotide and the deduced amino acid sequences from wild-type and putative c.7036+2T>A mutant are reported below of the respective protein diagrams. The nucleotide sequence is given in the upper line, and the amino acid translation (represented by single-letter code) is given below their respective codons. The nucleotide position in human *TG* mRNA is designated according to reference sequences (GenBank Accession No.: NM_003235). The A of the ATG of the initiator methionine codon is denoted nucleotide +1. The amino acid positions are numbered after subtracting the 19-amino-acid signal peptide. The resulting reading frame changes by exon 40 skipping are underlined. ⌜ indicates exon/exon boundaries and exon numbering is indicated.

this direction; NNSplice, FSplice, SPL, SPLM, MAXENT, MDD, MM and WMM tools showed a lower score for the cryptic 5′ ss than for the authentic 5′ ss. It is interesting to note that in contrast to the cryptic donor site upstream to c.745+1G>A mutation, the previously reported combined computational prediction algorithms in the cryptic donor site upstream to c.4159+3_+4delAT mutation showed a higher score relative to its corresponding wild type 5′ ss (Targovnik et al., 2012).

ESE elements are sequence tracts that are required for correct exon definition and inclusion (Blencowe, 2000). ESE sequences act as binding sites for SR proteins (SF2/ASF, SC35, SRp40 and SRp55) (Blencowe, 2000). ESE motifs were identified in the vicinity of the ss of constitutive and alternative exons (Wang et al., 2005). In this study we report the identification of ESE binding sites potentially recognizable by SF2/ASF, SC35, SRp40 and SRp55 proteins within the cryptic ss of exon 6 of the *TG* gene. We hypothesize that these ESE motifs may be involved in a splicing-regulatory pathway that is normally silent in the *TG* natural splicing environment, but when the wild type 5′ ss in intron 6 is disrupted by the c.745+1G>A mutation it can successfully influence cryptic ss usage. As expected, the c.745+1G>A mutation altered the ESE binding sites recognizable by SRp40 and SRp55 proteins in the authentic 5′ ss, according to ESEfinder program. In contrast, c.7036+2T>A mutation did not modify the recognition of ESE motifs by SR proteins.

The human *TG* preprotein monomer is composed of a 19-amino acids signal peptide followed by a 2749 residues polypeptide, organized in four structural and functional regions (Lee and Arvan, 2011; Lee et al., 2011). Region I contains 10 of the 11 *TG* type-1 repeats followed by a 247 residue hinge region, located between Type 1-10 and Type 2-1 domains. Region II contains three type-2 repeats and the 11th *TG* type-1 repeat, while region III contains five type-3 repeats (Figs. 6 and 7). Finally, the fourth region is located in the carboxy-terminal

domain and shows significant homology with the acetylcholinesterase (ACHE), named the ACHE-like domain. This domain is required for protein dimerization, which is essential for normal conformational maturation and intracellular transport of *TG* (Lee et al., 2009; Park and Arvan, 2004). ACHE-like domain may function as an intramolecular chaperone and as a molecular escort for Type 1, Type 2 and Type 3 *TG* domains (Lee et al., 2008). Newly synthesized *TG* is transported from the endoplasmic reticulum (ER) to the cell surface via the Golgi complex. It is interesting that all truncated forms detected in this study eliminate the ACHE-like domain. The c.745+1G>A mutation originates two putative truncated proteins of 200 amino acids when exon 6 is skipped, and 1142 amino acids when cryptic 5′ ss is activated (Fig. 6), whereas the c.7036+2T>A mutation results in a putative truncated protein of 2277 amino acids (Fig. 7). The functional consequences of the frameshift and subsequent premature stop codon, generated by missplicing of *TG* pre-mRNA, could be retention and accumulation of molecules devoid of the ACHE-like domain within the ER (Lee et al., 2008). ER quality control system prevents misfolded protein export from ER to Golgi and consequently fails to be transported to the site of thyroid hormone synthesis (Hammond and Helenius, 1995; Kim and Arvan, 1998; Kim et al., 2008). These gives rise to a distention of ER, abnormality called as ER-storage disease (ERSD) (Kim and Arvan, 1998). Misfolded proteins are degraded by the ER-associated degradation (ERAD) pathway (McCracken and Brodsky, 1996). On the other hand, the two truncated forms originated by c.745+1G>A mutation contain only the amino-terminal hormonogenic sites (the acceptor tyrosine 5 and the donor thyrosine 130). Whereas the aberrant *TG* generated by c.7036+2T>A mutation only lacks the carboxyl-terminal hormonogenic sites (positions: 2554 and 2774). Consequently, all mutant TG proteins have limited ability to generate thyroid hormone. It can be hypothesized that due to a structural distortion as a result of the truncated proteins, the ability to transfer an iodophenoxyl group from the donor site to the acceptor iodotyrosine is altered and therefore the formation of thyroid hormones is deficient.

From our study, we can draw four conclusions: First, functional analysis demonstrated that these two novel mutations inactivate the corresponding donor ss. In the case of the c.745+1G>A mutation, we were able to demonstrate the activation of a cryptic donor ss located in exon 6, in addition to the skipping of the entire exon 6, whereas, in the case of the c.7036+2T>A mutation, we evidenced only the skipping of the entire exon 40 from mature mRNA. Second, these results suggest the coexistence in our patient of three putative truncated proteins of 200, 1142 and 2277 amino acids. Third, we propose that the cause of congenital hypothyroidism in our patient is an interaction of two mechanisms, on the one hand the blocking of the *TG* transport in the secretory pathway with a consequent defective *TG* secretion to the apical luminal cavity of thyroid follicles due to the lack of ACHE-like domain, and on the other hand the deficient thyroid hormone synthesis because of the lack of the carboxyl-terminal hormonogenic sites. Finally, the present study provided qualitative support for the effect of both mutations on the pathophysiological cellular splicing events. The role of cryptic 5′ ss localized in the *TG* gene deserves intensive investigation, using *TG* as a model, in order to further understand the molecular mechanisms involved in splicing defects.

Acknowledgements

C.E. Citterio and C.M. Morales are research fellows of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

C.M. Rivolta and H.M. Targovnik are established investigators of the CONICET.

This study was supported by grants from the Universidad de Buenos Aires (20020100100594/2011 to CMR), CONICET (PIP 2012/112-201101-00091 to HMT), ANPCyT-FONCyT (PICT 2010/05-1130 to CMR, PICT 2012/05-1090 to HMT) and by FIS (PI10/00219 to RG-S).

References

Abdul-Hassan, I.A., AL-Ramahi, I.J., AL-Faisal, A.H.M., 2013. Detection of T.G. and TO genes compound mutations associated with thyroid carcinoma, toxic goiter and hypothyroidism in Iraqi patients. J. Med. Sci. 13, 676–683.

Agretti, P., De Marco, G., Di Cosmo, C., Ferrarini, E., Montanelli, L., Bagattini, B., et al., 2013. Congenital hypothyroidism caused by a novel homozygous mutation in the thyroglobulin gene. Eur. J. Pediatr. 172, 959–964.

Alzahrani, A.S., Baitei, E.Y., Zou, M., Shi, Y., 2006. Metastatic thyroid follicular carcinoma arising from congenital goiter due to a novel splice donor site mutation in the thyroglobulin gene. J. Clin. Endocrinol. Metab. 91, 740–746.

Bertaux, F., Noël, M., Malthiéry, Y., Fragu, P., 1991. Demonstration of a heterogeneous transcription pattern of thyroglobulin mRNA in human thyroid tissues. Biochem. Biophys. Res. Commun. 178, 586–592.

Bertaux, F., Noël, M., Lasmoles, F., Fragu, P., 1995. Identification of the exon structure and four alternative transcripts of the thyroglobulin-encoding gene. Gene 156, 297–301.

Blencowe, B.J., 2000. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. Trends Biochem. Sci. 25, 110–116.

Bonnet, C., Krieger, S., Vezain, M., Rousselin, A., Tournier, I., Martins, A., et al., 2008. Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene. J. Med. Genet. 45, 438–446.

Buratti, I., Chivers, M., Královičová, J., Romano, M., Baralle, M., Krainer, A.R., et al., 2007. Aberrant 5' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. Nucleic Acids Res. 37, 4250–4263.

Cangul, H., Boelaert, K., Dogan, M., Saglam, Y., Kendall, M., Barrett, T.G., et al., 2014. Novel truncating thyroglobulin gene mutations associated with congenital hypothyroidism. Endocrine 45, 206–212.

Caputo, M., Rivolta, C.M., Esperante, S.A., Gruñeiro-Papendieck, L., Chiesa, A., Pellizas, C.G., et al., 2007a. Congenital hypothyroidism with goitre caused by new mutations in the thyroglobulin gene. Clin. Endocrinol. (Oxf) 67, 351–357.

Caputo, M., Rivolta, C.M., Gutnisky, V.J., Gruñeiro-Papendieck, L., Chiesa, A., Medeiros-Neto, G., et al., 2007b. Recurrence of the p.R277X/p.R1511X compound heterozygous mutation in the thyroglobulin gene in unrelated families with congenital goiter and hypothyroidism: haplotype analysis using intragenic thyroglobulin polymorphisms. J. Endocrinol. 195, 167–177.

Caron, P., Moya, C.M., Malet, D., Gutnisky, V.J., Chabardes, B., Rivolta, C.M., et al., 2003. Compound heterozygous mutations in the thyroglobulin gene (1143delC and 6725G>A[R2223H]) resulting in fetal goitrous hypothyroidism. J. Clin. Endocrinol. Metab. 88, 3546–3553.

Citterio, C.E., Coutant, R., Rouleau, S., Miralles García, J.M., Gonzalez-Sarmiento, R., Rivolta, C.M., et al., 2011. A new compound heterozygous for c.886C>T/c.2206C>T [p.R277X/p.Q717X] mutations in the thyroglobulin gene as a cause of foetal goitrous hypothyroidism. Clin. Endocrinol. (Oxf) 74, 533–535.

Citterio, C.E., Machiavelli, G.A., Miras, M.B., Gruñeiro-Papendieck, L., Lachlan, K., Sobrero, G., et al., 2013a. New insights into thyroglobulin gene: molecular analysis of seven novel mutations associated with goiter and hypothyroidism. Mol. Cell. Endocrinol. 365, 277–291.

Citterio, C.E., Rossetti, L.C., Souchon, P.F., Morales, C., Thouvard-Viprey, M., Salmon-Musial, A.S., et al., 2013b. Novel mutational mechanism in the thyroglobulin gene: imperfect DNA inversion as a cause for hereditary hypothyroidism. Mol. Cell. Endocrinol. 381, 220–229.

Corral, J., Martín, C., Pérez, R., Sánchez, I., Mories, M.T., San Millan, J.L., et al., 1993. Thyroglobulin gene point mutation associated with non-endemic simple goitre. Lancet 341, 462–464.

Gamazon, E.R., Stranger, B.E., 2014. Genomics of alternative splicing: evolution, development and pathophysiology. Hum. Genet. 133, 679–687.

González-Sarmiento, R., Corral, J., Mories, M.T., Corrales, J.J., Miguel-Velado, E., Miralles-Garcia, J.M., 2001. Monoallelic deletion in the 5' region of the thyroglobulin gene as a cause of sporadic nonendemic simple goiter. Thyroid 11, 789–793.

Gutnisky, V.J., Moya, C.M., Rivolta, C.M., Domené, S., Varela, V., Toniolo, J.V., et al., 2004. Two distinct compound heterozygous constellation (R277X/IVS34-1G>C and R277X/R1511X) in the thyroglobulin (TG) gene in affected individuals of a Brazilian kindred with congenital goiter and defective TG synthesis. J. Clin. Endocrinol. Metab. 89, 646–657.

Hammond, C., Helenius, A., 1995. Quality control in the secretory pathway. Curr. Opin. Cell Biol. 7, 523–529.

Hermanns, P., Refetoff, S., Sriphrapradang, C., Pohlenz, J., Okamato, J., Slyper, L., et al., 2013. A clinically euthyroid child with a large goiter due to a thyroglobulin gene defect: clinical features and genetic studies. J. Pediatr. Endocrinol. Metab. 26, 119–123.

Hishinuma, A., Takamatsu, J., Ohyama, Y., Yokozawa, T., Kanno, Y., Kuma, K., et al., 1999. Two novel cysteine substitutions (C1263R and C1995S) of thyroglobulin cause a defect in intracellular transport of thyroglobulin in patients with congenital goiter and the variant type of adenomatous goiter. J. Clin. Endocrinol. Metab. 84, 1438–1444.

Hishinuma, A., Fukata, S., Kakudo, K., Murata, Y., Ieiri, T., 2005. High incidence of thyroid cancer in long-standing goiters with thyroglobulin mutations. Thyroid 15, 1079–1084.

Hishinuma, A., Fukata, S., Nishiyama, S., Nishi, Y., Oh-Ishi, M., Murata, Y., et al., 2006. Haplotype analysis reveals founder effects of thyroglobulin gene mutations C1058R and C1977S in Japan. J. Clin. Endocrinol. Metab. 91, 3100–3104.

Ieiri, T., Cochaux, P., Targovnik, H.M., Suzuki, M., Shimoda, S.-I., Perret, J., et al., 1991. A 3' splice site mutation in the thyroglobulin gene responsible for congenital goiter with hypothyroidism. J. Clin. Invest. 88, 1901–1905.
Kahara, T., Igarashi, N., Hishinuma, A., Nakanishi, Y., Uchiyama, A., Miwa, A., et al., 2012. Thyroglobulin gene mutation with cold nodule on thyroid scintigraphy. Case Reports Endocrinol. 2012, 280319.
Kanou, Y., Hishinuma, A., Tsunekawa, K., Seki, K., Mizuno, Y., Fujisawa, H., et al., 2007. Thyroglobulin gene mutations producing defective intracellular transport of thyroglobulin are associated with increased thyroidal type 2 iodothyronine deiodinase activity. J. Clin. Endocrinol. Metab. 92, 1451–1457.
Kim, P.S., Arvan, P., 1998. Endocrinopathies in the family of endoplasmic reticulum (ER) storage diseases: disorders of protein trafficking and the role of ER molecular chaperones. Endocr. Rev. 19, 173–202.
Kim, P.S., Lee, J., Jongsamak, P., Menon, S., Li, B., Hossain, S.A., et al., 2008. Defective protein folding and intracellular retention of thyroglobulin-R19K mutant as a cause of human congenital goiter. Mol. Endocrinol. 22, 477–484.
Kitanaka, S., Takeda, A., Sato, U., Miki, Y., Hishinuma, A., Ieiri, T., et al., 2006. A novel compound heterozygous mutation in the thyroglobulin gene resulting in congenital goitrous hypothyroidism with high serum triiodothyronine levels. J. Hum. Genet. 51, 379–382.
Lee, J., Di Jeso, B., Arvan, P., 2008. The cholinesterase-like domain of thyroglobulin functions as an intramolecular chaperone. J. Clin. Invest. 118, 2950–2958.
Lee, J., Wang, X., Di Jeso, B., Arvan, P., 2009. The cholinesterase-like domain, essential in thyroglobulin trafficking for thyroid hormone synthesis, is required for protein dimerization. J. Biol. Chem. 284, 12752–12761.
Lee, J., Arvan, P., 2011. Repeat motif-containing regions within thyroglobulin. J. Biol. Chem. 286, 26327–26333.
Lee, J., Di Jeso, B., Arvan, P., 2011. Maturation of thyroglobulin protein región I. J. Biol. Chem. 286, 33045–33052.
Liu, S., Zhang, S., Li, W., Zhang, A., Qi, F., Zheng, G., et al., 2012. Clinical and genetic analysis of a compound heterozygous mutation in the thyroglobulin gene in a Chinese twin family with congenital goiter and hypothyroidism. Twin Res. Hum. Genet. 15, 126–132.
Machiavelli, G.A., Caputo, M., Rivolta, C.M., Olcese, M.C., Gruñeiro-Papendieck, L., Chiesa, A., et al., 2010. Molecular analysis of congenital goiters with hypothyroidism caused by defective thyroglobulin synthesis. Identification of a novel c.7006C>T [p.R2317X] mutation and expression of minigenes containing nonsense mutations in exon 7. Clin. Endocrinol. (Oxf) 72, 112–121.
Mason, M.E., Dunn, A.D., Wortsman, J., Day, R.N., Day, K.H., Hoback, S.J., et al., 1995. Thyroids from siblings with Pendred's syndrome contain thyroglobulin messenger ribonucleic acid variants. J. Clin. Endocrinol. Metab. 80, 497–503.
McCracken, A.A., Brodsky, J.L., 1996. Assembly of ER-associated protein degradation *in vitro*: dependence on cytosol, calnexin, and ATP. J. Cell Biol. 132, 291–298.
Moya, C.M., Vallespin, E., Szkudlarek, A., Persani, L., Martin-Pena, M., Fugazzola, L., et al., 2011. A "customized" CGH-array thyroarray® identifies genetic defects in congenital hypothyroidism not detectable by PCR and sequencing. 35th Annual Meeting of the European Thyroid Association, Abstract OP66. Eur. Thyroid J. 0, 94.
Narumi, S., Muroya, K., Asakura, Y., Aachi, M., Hasegawa, T., 2011. Molecular basis of thyroid dyshormonogenesis: genetic screening in population-based Japanese patients. J. Clin. Endocrinol. Metab. 96, E1838–E1842.
Niu, D.M., Hsu, J.H., Chong, K.W., Huang, C.H., Lu, Y.H., Kao, C.H., et al., 2009. Six new mutations of the thyroglobulin gene discovered in taiwanese children presenting with thyroid dyshormonogenesis. J. Clin. Endocrinol. Metab. 94, 5045–5052.
Pardo, V., Rubio, I.G., Knobel, M., Aguiar-Oliveira, M.H., Santos, M.M., Gomes, S.A., et al., 2008. Phenotypic variation among four family members with congenital hypothyroidism caused by two distinct thyroglobulin gene mutations. Thyroid 18, 783–786.
Pardo, V., Vono-Toniolo, J., Rubio, I.G., Knobel, M., Possato, R.F., Targovnik, H.M., et al., 2009. The p.A2215D thyroglobulin gene mutation leads to deficient synthesis and secretion of the mutated protein and congenital hypothyroidism with wide phenotype variation. J. Clin. Endocrinol. Metab. 94, 2938–2944.
Park, Y.N., Arvan, P., 2004. The acetylcholinesterase homology region is essential for normal conformational maturation and secretion of thyroglobulin. J. Biol. Chem. 279, 17085–17089.
Pérez-Centeno, C., González-Sarmiento, R., Mories, M.T., Corrales, J.J., Miralles-García, J.M., 1996. Thyroglobulin exon 10 gene point mutation in a patient with endemic goiter. Thyroid 6, 423–427.
Peteiro-Gonzalez, D., Lee, J., Rodriguez-Fontan, J., Castro-Piedras, I., Cameselle-Teijeiro, J., Beiras, A., et al., 2010. New insights into thyroglobulin pathophysiology revealed by the study of a family with congenital goiter. J. Clin. Endocrinol. Metab. 95, 3522–3526.
Raef, H., Al-Rijjal, R., Al-Shehri, S., Zou, M., Al-Mana, H., Baitei, E.Y., et al., 2010. Biallelic p.R2223H mutation in the thyroglobulin gene causes thyroglobulin retention and severe hypothyroidism with subsequent development of thyroid carcinoma. J. Clin. Endocrinol. Metab. 95, 1000–1006.
Ricketts, M.H., Simons, M.J., Parma, J., Mercken, L., Dong, Q., Vassart, G., 1987. A nonsense mutation causes hereditary goitre in the Afrikander cattle and unmasks alternative splicing of thyroglobulin transcripts. Proc. Natl. Acad. Sci. USA 84, 3181–3184.
Rivolta, C.M., Moya, C.M., Gutnisky, V.J., Varela, V., Miralles-García, J.M., González-Sarmiento, R., et al., 2005. A new case of congenital goiter with hypothyroidism due to a homozygous p.R277X mutation in the exon 7 of the thyroglobulin gene: a mutational hot spot could explain the recurrence of this mutation. J. Clin. Endocrinol. Metab. 90, 3766–3770.
Roca, X., Sachidanandam, R., Krainer, A.R., 2003. Intrinsic differences between authentic and cryptic 5' splice sites. Nucleic Acids Res. 31, 6321–6333.
Sato, A., Abe, K., Yuzuriha, M., Fujii, S., Takahashi, N., Hojo, H., et al., 2014. A novel mutation in the thyroglobulin gene that causes goiter and dwarfism in Wistar Hannover GALAS rats. Mutat. Res. 762, 17–23.
Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. 15, 7155–7174.
Targovnik, H.M., Cochaux, P., Corach, D., Vassart, G., 1992. Identification of a minor Tg mRNA transcript in RNA from normal and goitrous thyroids. Mol. Cell. Endocrinol. 84, R23–R26.
Targovnik, H.M., Medeiros-Neto, G., Varela, V., Cochaux, P., Wajchenberg, B.L., Vassart, G., 1993. A nonsense mutation causes human hereditary congenital goiter with preferential production of a 171-nucleotide-deleted thyroglobulin ribonucleic acid messenger. J. Clin. Endocrinol. Metab. 77, 210–215.
Targovnik, H., Vono, J., Billerbeck, A.E.C., Cerrone, G.E., Varela, V., Mendive, F., et al., 1995. A 138-nucleotide deletion in the thyroglobulin ribonucleic acid messenger in a congenital goiter with defective thyroglobulin synthesis. J. Clin. Endocrinol. Metab. 80, 3356–3360.
Targovnik, H.M., Rivolta, C.M., Mendive, F.M., Moya, C.M., Medeiros-Neto, G., 2001. Congenital goiter with hypothyroidism caused by a 5′ splice site mutation in the thyroglobulin gene. Thyroid 11, 685–690.
Targovnik, H.M., Esperante, S.A., Rivolta, C.M., 2010a. Genetics and phenomics of hypothyroidism and goiter due to thyroglobulin mutations. Mol. Cell. Endocrinol. 322, 44–55.
Targovnik, H.M., Souchon, P.F., Machiavelli, G.A., Salmon-Musial, A.S., Mauran, P.L., Sulmont, V., et al., 2010b. Congenital goitre with hypothyroidism caused by a novel compound heterozygous mutations in the thyroglobulin gene. Clin. Endocrinol. (Oxf) 72, 716–718.
Targovnik, H.M., Citterio, C.E., Rivolta, C.M., 2011. Thyroglobulin Gene Mutations in Congenital Hypothyroidism. Horm. Res. Paediatr. 75, 311–321.
Targovnik, H.M., Edouard, T., Varela, V., Tauber, M., Citterio, C.E., González-Sarmiento, R., et al., 2012. Two Novel Mutations in the Thyroglobulin Gene as Cause of Congenital Hypothyroidism. Identification a Cryptic Donor Splice Site in the Exon 19. Mol. Cell. Endocrinol. 348, 313–321.
Targovnik, H.M., 2012. Thyroglobulin structure, function and biosynthesis. In: Braverman, L., Cooper, D. (Eds.), Werner and Ingbar's The Thyroid: A Fundamental and Clinical Text, Tenth ed. Lippincott Williams & Wilkins, Philadelphia, USA, pp. 74–92.
Tournier, I., Vezain, M., Martins, A., Charbonnier, F., Baert-Desurmont, S., Olschwang, S., et al., 2008. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum. Mutat. 29, 1412–1424.
van de Graaf, S.A.R., Ris-Stalpers, C., Veenboer, G.J.M., Cammenga, M., Santos, C., Targovnik, H.M., et al., 1999. A premature stopcodon in thyroglobulin mRNA results in familial goiter and moderate hypothyroidism. J. Clin. Endocrinol. Metab. 84, 2537–2542.
van de Graaf, S.A.R., Ris-Stalpers, C., Pauws, E., Mendive, F.M., Targovnik, H.M., de Vijlder, J.J.M., 2001. Up to date with human thyroglobulin. J. Endocrinol. 170, 307–321.
Wang, J., Smith, P.J., Krainer, A.R., Zhang, M.Q., 2005. Distribution of SR protein exonic splicing enhancer motifs in human protein-coding genes. Nucleic Acids Res. 33, 5053–5062.