OXFORD UNIVERSITY PRESS | Human Molecular Genetics

A FREQUENT OLIGOGENIC INVOLVEMENT IN CONGENITAL HYPOTHYROIDISM

Journal:	*Human Molecular Genetics*
Manuscript ID	HMG-2017-D-00181.R1
Manuscript Type:	2 General Article - UK Office
Date Submitted by the Author:	n/a
Complete List of Authors:	de Filippis, Tiziana; Istituto Auxologico Italiano - Istituto Scientifico San Luca, Lab of Endocrine and Metabolic Research Gelmini, Giulia; Istituto Auxologico Italiano Istituto di Ricovero e Cura a Carattere Scientifico, Lab of Endocrine and Metabolic Research paraboschi, Elvezia; Istituto Clinico Humanitas, Genetics Vigone, Maria Cristina; Ospedale San Raffaele, Pediatrics Di Frenna, Marianna; Ospedale San Raffaele, Pediatrics Marelli, Federica; Istituto Auxologico Italiano Istituto di Ricovero e Cura a Carattere Scientifico, Lab of Endocrine and Metabolic Research Bonomi, Marco; Istituto Auxologico Italiano Istituto di Ricovero e Cura a Carattere Scientifico, Lab of Endocrine and Metabolic Research Cassio, Alessandra; Universita degli Studi di Bologna, Pediatrics Larizza, Daniela; Ospedale San Matteo IRCCS, Peadiatrics Moro, Mirella; Istituto Auxologico Italiano Istituto di Ricovero e Cura a Carattere Scientifico, Divisione of Endocrinology and Metabolic Diseases Radetti, Giorgio; Ospedale Generale Regionale di Bolzano, Pediatrics Salerno, Maria; University "Federico II" of Naples, Pediatric Endocrinology Unit, Department of Translational Medical Sciences Ardissino, Diego; Azienda Ospedaliero-Universitaria di Parma, Division of Cardiology Weber, Giovanna; Vita-Salute University, San Raffaele Scientific Institute, Department of Pediatrics Gentilini, Davide; IRCCS Istituto Auxologico Italiano, Laboratory of Molecular Biology Guizzardi, Fabiana; Istituto Auxologico Italiano IRCCS, Unità di Medicina Generale ad indirizzo Endocrino-Metabolico e Lab. di Ricerche Endocrino-Metaboliche Duga, Stefano; Università degli Studi di Milano, Biologia e Genetica per le Scienze Mediche Persani, Luca; Institute of Endocrine Sciences, University of Milan, IRCCS Istituto Auxologico Italiano
Key Words:	Congenital hypothyroidism, Thyroid, Next Generation Sequencing, Thyroid dysgenesis, Goiter

A FREQUENT OLIGOGENIC INVOLVEMENT IN CONGENITAL HYPOTHYROIDISM

Tiziana de Filippis[§1], Giulia Gelmini[§1], Elvezia Paraboschi[§2,3], Maria Cristina Vigone[4], Marianna Di Frenna[4], Federica Marelli[1], Marco Bonomi[1,5], Alessandra Cassio[6], Daniela Larizza[7], Mirella Moro[1], Giorgio Radetti[8], Mariacarolina Salerno[9], Diego Ardissino[10], Giovanna Weber[4], Davide Gentilini[1], Fabiana Guizzardi[1], Stefano Duga[2,3], Luca Persani [*1,5]

[1]Division of Endocrine and Metabolic Diseases & Labs of Endocrine and Metabolic Research or Molecular Biology, IRCCS Istituto Auxologico Italiano, Milan; [2]Department of Biomedical Sciences, Humanitas University, Via Manzoni 113, 20089 Rozzano, Milan, Italy; [3]Humanitas Clinical and Research Center, Via Manzoni 56, 20089 Rozzano, Milan, Italy; [4]San Raffaele Hospital and Vita-Salute University, Milan; [5]Department of Clinical Sciences and Community Health (DISCCO), University of Milan; [6]Department of Pediatrics, University of Bologna; [7]Department of Pediatrics, IRCCS Policlinico San Matteo, University of Pavia; [8]Bolzano Hospital, Bolzano; [9]Department of Pediatrics, University of Naples, Naples; [10]Cardiology Department, Parma University Hospital, Parma, Italy

[§]These authors contributed equally to this work and should be considered as first author.

Correspondence to:
Luca Persani, MD PhD
Division of Endocrine Metabolic Diseases, San Luca Hospital
Piazzale Brescia 20 – 20149 Milan – Italy
Phone: +390261911-2738; fax: -2777
Email: luca.persani@unimi.it

Abstract

Congenital Hypothyroidism (CH), the most frequent form of preventable mental retardation, is predicted to have a relevant genetic origin. However, CH is frequently reported to be sporadic and candidate gene variations were found in <10% of the investigated patients. Here, we characterize the involvement of 11 candidate genes through a systematic Next Generation Sequencing (NGS) analysis. The NGS was performed in 177 unrelated CH patients (94 gland-in-situ; 83 dysgenesis) and in 3,538 control subjects. Non-synonymous or splicing rare variants (MAF<0.01) were accepted, and their functional impact was predicted by a comprehensive bioinformatic approach and co-segregation studies. The frequency of variations in cases and controls was extended to 18 CH-unrelated genes.

At least one rare variant was accepted in 103/177 patients. Monogenic recessive forms of the disease were found in 5 cases, but oligogenic involvement was detected in 39 patients. The 167 variations were found to affect all genes independently of the CH phenotype. These findings were replicated in an independent cohort of additional 145 CH cases. When compared to 3,538 controls, the CH population was significantly enriched with disrupting variants in the candidate genes ($p=5.5*10^{-7}$), but not with rare variations in CH-unrelated genes. Co-segregation studies of the hypothyroid phenotype with multiple gene variants in several pedigrees confirmed the potential oligogenic origin of CH.

The systematic NGS approach reveals the frequent combination of rare variations in morphogenetic or functional candidate genes in CH patients independently of phenotype. The oligogenic origin represents a suitable explanation for the frequent sporadic CH occurrence.

Introduction

Congenital hypothyroidism (CH) is the most common endocrine developmental disorder and preventable cause of mental retardation (1). A tendency toward an increased CH incidence or detection has been described in various socio-economically developed countries (2-4). Both functional and developmental defects can account for CH but its etiology is still poorly understood (5,6).

It is classically described to be sporadic but several findings in humans and experimental models support a relevant genetic origin. CH is currently considered as a puzzle of monogenic diseases and the investigations on its pathogenesis were driven by phenotype and performed by Sanger sequencing (5-10). Indeed, pathogenic variations in candidate genes have been found in <10% of the cases (5), but genetic studies targeting specific phenotypes or particular ethnic groups yielded higher mutation detection rates (11-14). On these bases, CH can be classified as a disease with a strong genetic component, but with a largely missing heritability (15).

Next generation sequencing (NGS) allows the simultaneous and systematic analysis of candidate genes in unselected disease populations, likely opening novel perspectives in the diagnosis and classification of several diseases (15, 16), but no data are available on the application of this strategy in a large number of CH patients.

Here, we systematically analyzed by NGS a panel of 11 candidate genes in 177 unrelated patients with CH, the identified variations were pathogenically classified by a comprehensive bioinformatic approach and genotype-phenotype co-segregation studies in families. The frequency of accepted gene variants (AVs) was then checked in a large control sample of 3,538 subjects from the same geographic area and in an independent CH cohort (n=145).

Results

The NGS analysis of the 11 candidate genes in 177 unrelated CH patients identified 167 allelic variants (116 AVs with MAF <0.01 and 51 novel AVs; 23 AVs were found more than once in

homozygosity or in combination with other variations)(Table S2). At least 1 variant was identified in 103/177 patients (58.2%). Variations were found in all CH categories, but patients with gland-in-situ (GIS) and hypoplasia tended to be more frequently affected (Figure 1A). A significant number of patients have ≥2 AVs in the same or in different candidate genes (44; 24.8%) (Figure 1B). Five of these cases had biallelic (homozygous or compound heterozygous) gene variants with a phenotype corresponding to a recessive mode of inheritance: one hypoplasia associated with a compound heterozygosity in *FOXE1* (p.S152W/p.Y192LfsX37), another hypoplasia with a compound heterozygosity in *TSHR* (p.I640L/p.R310H), and 3 GIS: one with a homozygous *TG* variant (p.T411HfsX80), one with a compound heterozygosity in *DUOX2* (p.R516H/p.D870A), and one with compound heterozygosity in *TPO* gene (p.A489T/p.E799K). These five patients with monogenic disease were excluded from further analyses. The list of identified variants is reported in Table S2, and their distribution in CH carriers is reported in Figure S1. A variable number of AVs were detected in the 11 candidate genes and most of them were missense variants (Figure 2A). The genes affected with the higher number of AVs were *DUOX2, TG* and *TSHR*. The 8 *PAX8* and 2 *DUOXA2* AVs were found exclusively in 10 patients with GIS, but all other genes, independently of their role in thyroid morphogenesis or function, were associated either with the dysgenesis or GIS phenotype (Figure 2B). The combinations of gene variants appear to occur casually (Figure S1), and biallelic variations in two functional genes, such as *TG* or *SLC26A4,* were associated with thyroid dysgenesis in CH cases #7 (athyreosis) and #14 (ectopy) that are also affected with extra-thyroid manifestations (Figure S1 and Table S3).

The distribution of extra-thyroid manifestations among the patients with or without AVs is reported in Table S3. As expected, choreoathetosis and/or hypotonia were found in 3 carriers of *NKX2-1* AVs and heart malformations in 2 carriers of *JAG1* defects; however, *SLC26A4* variations were surprisingly prevalent in CH patients with concomitant heart defects (7/20 cases).

The frequency of potential pathogenic variations within the same 11 candidate genes was assessed in a cohort of Italian individuals. The analysis of the principal components revealed that the case

and control populations were comparable (data not shown). When considering all the AVs, a relevant number of control subjects (43.5%) resulted to be carrier of >1 rare variant, however only a smaller fraction (11.7%) was carrier of multiple AVs, if compared to the CH group (23.2%). The overall burden of variants resulted statistically different between cases and controls ($p = 1.2*10^{-6}$), consistent with an enrichment of rare variations within the CH population (Figure 3). Moreover, a significantly higher number of CH patients resulted to be carrier of at least one disruptive AV (frameshift, nonsense and splicing variants, or missense variants previously reported to be loss-of-function in functional tests or familial linkage, or predicted to be deleterious by at least 5 out of the 7 used algorithms) in comparison with controls (32.9 vs 19.1%; $p = 5.5*10^{-7}$, Figure 3).

To confirm the specificity of these results, we performed a parallel analysis of 18 CH-unrelated genes in which again the distribution of rare variants was compared between cases and controls. As expected in this analysis, no significant difference in the mutation burden between the two groups was detected ($p = 0.1982$, Figure 3).

The analysis of the AVs in nine familial settings reveals variable combinations of inherited and de novo variations in the CH probands that are consistent with an oligogenic model of the disease. inherited multiple rare variants in distinct genes from the two parents, that were either affected with mild forms of non-autoimmune hypothyroidism or apparently unaffected (see details in Figure 4).

Finally, the NGS analysis was replicated in an independent cohort of 145 CH patients: we found 91 carriers (62.7%) of ≥1 AVs in the 11 genes. At least 2 AVs were detected in 41/145 CH patients (28.3%); since 3/41 had monogenic defects, oligogenic defects were found in 38/145 patients (26.2% of total).

Discussion

The systematic NGS analysis of 11 known candidate genes provides evidence that inheritable variations with a proven or predicted loss-of-function effect can be found in a large portion of CH patients. According to what was anticipated in a mouse model (17), we show an oligogenic

involvement in 22% of total cases, with a likelihood of oligogenicity among mutated patients of about 42.7%. Importantly, these results were replicated in an independent CH cohort and CH candidate gene variations were found also in the control population, but with a significantly lower prevalence. Thus, our results indicate that the pathogenesis of CH can frequently be due to the sum of rare (MAF <0.01) alleles (15) that generate minor functional impairments and are not expected to be associated individually with an overt phenotype. This oligogenic origin of CH is confirmed by the co-segregation of variants with CH or non-autoimmune hypothyroid state in several families. Therefore, our findings give a significant contribution to the missing heritability of CH, and may constitute a suitable explanation for the variable modes of inheritance described in several pedigrees (18), as well as for the minor glandular defects detected in several relatives of CH patients (19).

So far, monogenic defects had been described in rare familial forms of CH, mainly associated with biallelic LOF mutations in functional genes (5-7, 9, 13, 20-22). Instead, heterozygous mutations in thyroid transcription factor genes were associated with thyroid defects with variable penetrance and expressivity, but several phenotype driven studies with previous screening methods (such as denaturing gel gradient electrophoresis or Sanger sequencing) in large CH cohorts reported a poor rate of positive results (5-7, 23-25). Our approach allowed the identification of classic recessive monogenic CH forms in five cases, but unexpected variations in genes that would have been excluded a priori from screening on the basis of the clinical phenotype. The AVs involve all the examined genes with different frequencies and independently if they were previously linked to dysgenetic or functional defects, thus justifying the previous low rate of variant detection following the classic phenotypical approach. It is worth noting that we identified biallelic mutations of *FOXE1* in a CH case affected with thyroid hypoplasia but not with the classic manifestations of Bamforth-Lazarus syndrome (athyreosis, cleft palate and spiky hair), and monoallelic *FOXE1* AVs were indeed found in GIS cases; in both of this conditions the analysis of *FOXE1* would have been excluded a priori. Furthermore, the association of biallelic AVs in functional genes such as *SLC26A4* (CH14 in Table S2) and *TG* (CH7 in Table S2) with thyroid dysgenesis is particularly

surprising. A previous study (26) proposed a "toxic" effect produced by *SLC26A4* inactivation on thyroid follicles as a possible explanation for this association. However, the co-existence of extra-thyroid defects, such as Arnold Chiari and severe short stature, in cases #14 and #7 suggests the possible combination with still uncovered defects in developmental genes. Moreover, we found *SLC26A4* variants equally distributed among cases with thyroid dysgenesis or GIS, and frequently associated with mild congenital heart defects. This latter finding is particularly intriguing because *SLC26A4* transcripts are expressed also in the heart tissue under the control of steroid hormones (27).

As recently reported (28, 29), variations in candidate genes were indeed detected more frequently by NGS than previously observed. But these studies were conducted in small series of familial CH with gland-in-situ (<49 subjects from 34 ethnically diverse families) of variable geographic origin. At variance, our study has been performed in a larger cohort of unrelated CH patients, including either dysgenetic and functional thyroid defects thus representing the general CH population.

An added value of our study is the comparison of the frequency of gene variations in a large Italian control population. This analysis reveals that the CH population is indeed significantly enriched with rare/low frequency alleles in the 11 CH-related genes, and the frequency of multiple gene involvement is 2- to 4-fold higher than in the control population. Importantly, the distribution difference between cases and controls: a) is higher when only disruptive variants in the CH related genes are taken into account, and b) disappears when the variations in 18 CH-unrelated genes are considered. It is then conceivable that variants with a modest functional impairment can produce a negligible effect on thyroid function when expressed alone, but the sum of minor alleles, even acting at different levels (thyroid morphogenesis or hormonogenesis), can justify the birth of a child with CH in families with a history of minor thyroid defects or without any previously recognized thyroid disease (Figure 4). Our findings extend the relevance of the genetic components of CH, supporting the hypothesis that sporadic CH may be the result of the combination of multiple defects in different CH genes. This oligogenic model may indeed constitute another suitable explanation for

the variable expressivity and penetrance of genetic defects previously reported in several CH familial settings (5, 6, 9, 18, 19, 23), and well fits the previously proposed animal model of a multifactorial pathogenesis of CH (17). On this line, we recently reported the involvement of *JAG1* loss-of-function variants in CH pathogenesis (10). In this context, *JAG1* may act as a gene modifier contributing to uncover a mild functional defect due to the concomitant presence of rare alleles in "specific thyroid genes". This is outlined in the pedigrees of CH cases #11 and #20, and may explain the first association of *NKX2-1* mutations with thyroid ectopy in case #12 (30). Therefore, we propose that beside rare monogenic forms, CH may more frequently arise in sporadic cases as a multifactorial disease with a relevant genetic predisposition. Indeed, environmental factors such as endocrine disruptors (or iodine deficiency) could contribute to endocrine dysfunction by influencing gene expression (31,32) and generating a more profound phenotype in carriers of rare genetic variants.

The NGS approach was predicted to become a useful tool for a more precise diagnosis and classification of affected patients (6). Its application to congenital hypothyroidism may overcome the current phenotypical classification based on cumbersome clinical and biochemical investigations (scintigraphy, perchlorate discharge test, ultrasound, biochemical thyroid function tests) and offer the possibility to increase the number of positive cases and predict the risk of CH or other forms of hypothyroidism in affected families. In the present experience, we could identify predisposing variants in the majority of CH babies and the study of their segregation in families revealed several relatives with a previously unknown hypothyroidism (Figure 4). The extension of the NGS panel to other candidates will indeed increase the performance of the NGS approach.

In conclusion, the first systematic NGS analysis performed in a large CH population reveals a frequent oligogenic origin of CH that may represent a suitable explanation for the missing heritability of CH and the prevalent sporadic presentation of the disease. Further studies are needed to prove if these rare alleles may contribute a significant portion of the functional and structural

thyroid defects in the general population that are not justified by autoimmunity or iodide deficiency (33,34).

Materials and Methods

Subjects

The patients were enrolled in several Italian referral centers within a specific research protocol (RF-2010-2309484) that was approved by the Ethics Committees of the involved institutions. Informed consent was obtained from the parents of each CH patient. The analysis was applied to the first 177 eligible subjects enrolled after the approval of the protocol.

The inclusion criteria were: positive neonatal TSH screening (dry blood spot TSH ≥10 mU/L) with diagnosis of primary CH confirmed by serum thyroid function tests at 1-3 weeks of age (serum TSH range: 9.2-951 mU/L; normal range: 0.4-7.5; serum FT4: 0.1-1.69 ng/dL; normal range: 1.5-2.4).

The exclusion criteria were: i) high urinary iodine; ii) anti-TSH receptor antibody in serum at neonatal age; iii) premature birth (<35 weeks of gestation).

All patients had a phenotypical classification by neck ultrasound and/or scintiscan. A thyroid dysgenesis was described in 83 CH patients, whereas a GIS of normal or enlarged size was described in 94 cases (Table 1). Additional information on possible existence of thyroid disease in members of the family or associated malformations/diseases was collected in all cases. In case of positive genetic investigations, we extended whenever possible the biochemical and genetic studies to the available first-degree relatives. The data describing the clinical and biochemical characteristics of the 177 unrelated patients is reported in Table 1. The genetic investigations were replicated in an independent cohort of 145 unrelated CH patients (40.7% dysgenesis; 59.3% GIS).

The control population (n= 3,538) was recruited within the "Atherosclerosis, Thrombosis, and Vascular Biology Italian Study Group", as previously described (35).

Genetic analyses

The genomic DNA of each patient was extracted from peripheral blood lymphocytes using Gene Catcher gDNA 96x10 ml Automated Blood kit (Invitrogen, Life TechnologiesTM). Genetic analyses were performed by NGS of a panel of 29 genes, including 11 CH candidate genes (*NKX2-1, PAX8, FOXE1, GLIS3, JAG1, TSHR, SLC26A4, TG, TPO, DUOX2, DUOXA2*)(Table S1) and 18 CH-unrelated genes. The TruSeq Custom Amplicon assay (Illumina, San Diego, CA) was designed by the GenomeStudio software, having 681 amplicons covering exons with 25 bases of the flanking introns; the total coverage of the target genes by the designed amplicons was 80%. All regions not sequenced were recovered by Nextera® DNA Library Preparation kit (Illumina, San Diego, CA). The TruSeq NGS libraries were prepared according to the manufacturer's instructions (Truseq® Custom Amplicon Library Preparation, Illumina, San Diego,CA). Pooled libraries were sequenced on MiSeq Reagent kit v.3 on an Illumina MiSeq sequencer (Illumina, San Diego,CA). Basic data analysis was performed according to the default parameters of the Illumina's MiSeq Reporter software. For each subject, a .vcf file containing variant calls was generated, further reviewed, and filtered. For every equivocal call in the .vcf files, visual inspection of the mapped data was performed using the Integrated Genomics Viewer 2.3 software (IGV; Broad Institute, Cambridge, MA, USA). All variants of interest were verified by conventional dideoxy sequencing using BigDye® Terminator v.3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA) on a 3100 DNA Analyzer from Applied Biosystems (Foster City, CA). Whole-exome sequencing (WES) of the control cohort was performed at the Broad Institute (Boston, MA). Sequencing, exome capture methods, variant annotation, and data processing of the samples were previously described (36). Since data in the case and control cohorts were generated by means of different sequencing platforms, the same quality filters for strand bias, coverage, and mapping quality were applied.

Bioinformatics and statistical analyses

Only nonsense, frame-shift, splice-site (including the first 6 and the last 3 nucleotides of each intron), and missense variants with a minor allele frequency (MAF) <0.01 were considered in the

analysis (accepted variant, AV), i.e. rare variants (15). Computational analyses were carried out to predict the possible pathogenic role of missense variants, through the use of the dbNSFP database (37). In detail, we considered as disruptive the nonsense, frame-shift, splice-site mutations and the non-synonymous missense variants annotated as deleterious by at least 5 out of 7 algorithms of the dbNSFP database, i.e.: SIFT, PolyPhen2, MutationTaster, MutationAssessor, LRT (Likelihood Ratio Test), and FATHMM (Functional Analysis Through Hidden Markov Model). And, for the intronic variants close to the splicing sites (-3/+6 nucleotides from the boundary) we used: NetGene2v.2.4**, ESEfinder2.0, BDGP (see Table S2).

The frequency and the functional annotation of the identified variants were checked in public and licensed databases (Ensembl, UCSC Genome browser, 1000 Genome project, ExAC Browser, NCBI, HGMD professional).

Principal component analysis was performed on case and control populations using the SmartPCA program contained in the EIGENSOFT package (38,39).

Frequency and distribution of the variants in cases and controls were compared by means of Fisher's exact test and R software (40).

ACKNOWLEDGMENTS

The Authors wish to thank the patients and their families for the participation to the study.

Partially supported by the Italian Ministry of Health, Rome, Italy (grant: RF-2010-2309484).

CONFLICT OF INTEREST: The authors have nothing to disclose.

REFERENCES

1. Léger, J., Olivieri, A., Donaldson, M., Torresani, T., Krude, H., van Vliet, G., Polak, M., Butler, G., ESPE-PES-SLEP-JSPE-APEG-APPES-ISPAE., Congenital Hypothyroidism Consensus Conference Group. (2014) European Society for Paediatric Endocrinology consensus guidelines on screening, diagnosis, and management of congenital hypothyroidism. *J. Clin. Endocrinol.Metab*. **99**, 363-384.

2. Corbetta, C., Weber, G., Cortinovis, F., Calebiro, D., Passoni, A., Vigone, M.C., Beck-Peccoz, P., Chiumello, G., Persani, L. (2009) A 7-year experience with low blood TSH cutoff levels for neonatal screening reveals an unsuspected frequency of congenital hypothyroidism (CH). *Clin. Endocrinol. (Oxf.)* **71**, 739-745.

3. Olivieri, A., Corbetta, C., Weber, G., Vigone, M.C., Fazzini, C., Medda, E., Italian Study Group for Congenital Hypothyroidism. (2013) Congenital hypothyroidism due to defects of thyroid development and mild increase of TSH at screening: data from the Italian National Registry of infants with congenital hypothyroidism. *J. Clin. Endocrinol. Metab.* **98**, 1403-1408.

4. Rapaport, R. (2010) Congenital hypothyroidism: an evolving common clinical conundrum. *J. Clin. Endocrinol. Metab.* **95**, 4223-4225.

5. De Felice, M., Di Lauro, R. (2004) Thyroid development and its disorders: genetics and molecular mechanisms. *Endocr. Rev.* **25**, 722-746.

6. Park, S.M., Chatterjee, V.K. (2005) Genetics of congenital hypothyroidism. *J. Med. Genet.* **42**, 379-389.

7. Zamproni, I., Grasberger, H., Cortinovis, F., Vigone, M.C., Chiumello, G., Mora, S., Onigata, K., Fugazzola, L., Refetoff, S., Persani, L., et al. (2008) Biallelic inactivation of the dual oxidase maturation factor 2 (DUOXA2) gene as a novel cause of congenital hypothyroidism. *J. Clin. Endocrinol. Metab.* **93**, 605-610.

8. Dimitri, P., Habeb, A.M., Gurbuz, F., Millward, A., Wallis, S., Moussa, K., Akcay, T., Taha, D., Hogue, J., Slavotinek, A., et al. (2015) Expanding the Clinical Spectrum Associated With GLIS3 Mutations. *J. Clin. Endocrinol. Metab*. **100**, E1362-1369.

9. Moreno, J.C., Klootwijk, W., van Toor, H., Pinto, G., D'Alessandro, M., Lèger, A., Goudie, D., Polak, M., Grüters, A., Visser, T.J. (2008) Mutations in the iodotyrosine deiodinase gene and hypothyroidism. *N. Engl. J. Med*. **358**, 1811-1818.

10. de Filippis, T., Marelli, F., Nebbia, G., Porazzi, P., Corbetta, S., Fugazzola, L., Gastaldi, R., Vigone, M.C., Biffanti, R., Frizziero, D., et al. (2016) JAG1 Loss-of-function variations as a novel predisposing event in the pathogenesis of congenital thyroid defects. *J. Clin. Endocrinol. Metab.* **101**, 861-870.

11. Thorwarth, A., Schnittert-Hübener, S., Schrumpf, P., Müller, I., Jyrch, S., Dame, C., Biebermann, H., Kleinau, G., Katchanov, J., Schuelke, M., et al. (2014) Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum. *J. Med. Genet.* **51**, 375-387.

12. Muzza, M., Rabbiosi, S., Vigone, M.C., Zamproni, I., Cirello, V., Maffini, M.A., Maruca, K., Schoenmakers, N., Beccaria, L., Gallo, F., Park, S.M., et al. (2014) The clinical and molecular characterization of patients with dyshormonogenic congenital hypothyroidism reveals specific diagnostic clues for DUOX2 defects. *J. Clin. Endocrinol. Metab.* **99**, E544-553.

13. Persani, L., Calebiro, D., Cordella, D., Weber, G., Gelmini, G., Libri, D., de Filippis, T., Bonomi, M.. (2010) Genetics and phenomics of hypothyroidism due to TSH resistance. *Mol. Cell. Endocrinol.* **322**, 72-82.

14. Tenenbaum-Rakover, Y., Grasberger, H., Mamanasiri, S., Ringkananont, U., Montanelli, L., Barkoff, M.S., Dahood, A.M., Refetoff, S.. (2009) Loss-of-function mutations in the thyrotropin

receptor gene as a major determinant of hyperthyrotropinemia in a consanguineous community. *J. Clin. Endocrinol. Metab.* **94**, 1706-1712.

15. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009) Finding the missing heritability of complex diseases. *Nature.* **461**, 747-753.

16. Lu, J.T., Campeau, P.M., Lee, B.H.. (2014) Genotype--phenotype correlation-promiscuity in the era of next-generation sequencing. *N Engl J Med.* **371**, 593-596.

17. Amendola, E., De Luca, P., Macchia, P.E., Terracciano, D., Rosica, A., Chiappetta, G., Kimura, S., Mansouri, A., Affuso, A., Arra, C., et al. (2005) A mouse model demonstrates a multigenic origin of congenital hypothyroidism. *Endocrinology.* **146**, 5038-5047.

18. Castanet, M., Sura-Trueba, S., Chauty, A., Carré, A., de Roux, N., Heath, S., Léger, J., Lyonnet, S., Czernichow, P., Polak, M. (2005) Linkage and mutational analysis of familial thyroid dysgenesis demonstrate genetic heterogeneity implicating novel genes. *Eur. J. Hum. Genet.* **13**, 232-239.

19. Léger, J., Marinovic, D., Garel, C., Bonaïti-Pellié, C., Polak, M., Czernichow, P. (2002) Thyroid developmental anomalies in first degree relatives of children with congenital hypothyroidism. *J. Clin. Endocrinol. Metab.* **87**, 575-580.

20. Abramowicz, M.J., Targovnik, H.M., Varela, V., Cochaux, P., Krawiec, L., Pisarev, M.A., Propato. F.V., Juvenal, G., Chester, H.A., Vassart, G. (1992) Identification of a mutation in the coding sequence of the human thyroid peroxidase gene causing congenital goiter. *J. Clin. Invest.* **90**, 1200-1204.

21. Sunthornthepvarakul, T., Gottschalk, M.E., Hayashi, Y., Refetoff, S. (1995) Brief report: resistance to thyrotropin caused by mutations in the thyrotropin-receptor gene. *N. Engl. J. Med.* **332**,155-160.

22. Fugazzola, L., Mannavola, D., Vigone, M.C., Cirello, V., Weber, G., Beck-Peccoz, P., Persani, L. (2005) Total iodide organification defect: clinical and molecular characterization of an Italian family. *Thyroid.* **15**,1085-1088.

23. Macchia, P.E., Lapi, P., Krude, H., Pirro, M.T., Missero, C., Chiovato, L., Souabni, A., Baserga, M., Tassi, V., Pinchera, A., et al. (1998) PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. *Nat. Genet.* **19**, 83-86.

24. Lapi, P., Macchia, P.E., Chiovato, L., Biffali, E., Moschini, L., Larizza, D., Baserga, M., Pinchera, A., Fenzi, G., Di Lauro, R. (1997) Mutations in the gene encoding thyroid transcription factor-1 (TTF-1) are not a frequent cause of congenital hypothyroidism (CH) with thyroid dysgenesis. *Thyroid.* **7**, 383-387.

25. Taji, E., Biebermann, H., Límanová, Z., Hníková, O., Zikmund, J., Dame, C., Grüters, A., Lebl,J., Krude, H. (2007) Screening for mutations in transcription factors in a Czech cohort of 170 patients with congenital and early-onset hypothyroidism: identification of a novel PAX8 mutation in dominantly inherited early-onset non-autoimmune hypothyroidism. *Eur. J. Endocrinol.* **156**, 521-529.

26. Kühnen, P., Turan, S., Fröhler, S., Güran, T., Abali, S., Biebermann, H., Bereket, A., Grüters, A., Chen, W., Krude, H. (2014) Identification of PENDRIN (SLC26A4) mutations in patients with congenital hypothyroidism and "apparent" thyroid dysgenesis. *J. Clin. Endocrinol. Metab.* **99**, E169-176.

27. Pelzl, L., Pakladok, T., Pathare, G., Fakhri, H., Michael, D., Wagner, C.A., Paulmichl, M., Lang, F. (2012) DOCA sensitive pendrin expression in kidney, heart, lung and thyroid tissues. *Cell. Physiol. Biochem.* **30**, 1491-501.

28. Löf, C., Patyra, K., Kuulasmaa, T., Vangipurapu, J., Undeutsch, H., Jäschke, H., Pajunen, T., Kero, A., Krude, H., Biebermann, H., et al. (2016) Detection of novel gene variants associated with congenital hypothyroidism in a Finnish patient cohort. *Thyroid.* **26**, 1215-1224.

29. Nicholas, A.K., Serra, E.G., Cangul, H., Alyaarubi, S., Ullah, I., Schoenmakers, E., Deeb, A., Habeb, A.M., Almaghamsi, M., Peters, C., et al. (2016) Comprehensive screening of eight known causative genes in congenital hypothyroidism with gland-in-situ. *J. Clin. Endocrinol. Metab.* **101**, 4521-4531.

30. de Filippis, T., Marelli, F., Vigone, M.C., Di Frenna, M., Weber, G., Persani, L. *(*2014) Novel NKX2-1 frameshift mutations in patients with atypical phenotypes of the brain-lung-thyroid syndrome. *Eur. Thyroid J.* **3,** 227-233.

31. Gentilcore, D., Porreca, I., Rizzo, F., Ganbaatar, E., Carchia, E., Mallardo, M., De Felice, M., Ambrosino, C. (2013) Bisphenol A interferes with thyroid specific gene expression. *Toxicology.* **304**, 21-31.

32. Porreca, I., Ulloa-Severino, L., Almeida, P., Cuomo, D., Nardone, A., Falco, G., Mallardo, M., Ambrosino, C. (2017) Molecular targets of developmental exposure to bisphenol A in diabesity: a focus on endoderm-derived organs. *Obes. Rev.* **18**, 99-108.

33. Tunbridge, W.M., Evered, D.C., Hall, R., Appleton, D., Brewis, M., Clark, F., Evans, J.G., Young, E., Bird, T., Smith, P.A.. (1977) The spectrum of thyroid disease in a community: the Whickham survey. *Clin. Endocrinol. (Oxf).* **7**, 481–493.

34. Valdes, S., Maldonado-Araque, C., Lago-Sampedro, A., Lillo, J.A., Garcia-Fuentes, E., Perez-Valero, V., Gutierrez-Repiso, C., Ocon-Sanchez, P., Goday, A., Urrutia, I., et al. (2017) Population-based national prevalence of thyroid dysfunction in Spain and associated factors. Di@bet.es study. *Thyroid.* **27**, 156-166.

35. Atherosclerosis, Thrombosis, and Vascular Biology Italian Study Group. (2003) No evidence of association between prothrombotic gene polymorphisms and the development of acute myocardial infarction at a young age. *Circulation.* **107**, 1117-1122.

36. Do, R., Stitziel, N.O., Won, H.H., Jørgensen, A.B., Duga, S., Merlini, P.A., Kiezun, A., Farrall, M., Goel, A., Zuk, O., et al. (2015) Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature.* **518**, 102-106.

37. Liu, X., Jian, X., Boerwinkle, E. (2011) dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum. Mutat.* **32**, 894-899.

38. Price, A.L., Patterson, N.J., Plenge, R.M,, Weinblatt, M.E., Shadick, N.A., Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904-909.

39. Patterson, N., Price, A.L., Reich, D.. (2006) Population structure and eigenanalysis. *PLoS Genet.* **2**, e190.

40. R Development Core Team. (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, (ISBN 3-900051-07-0).

LEGENDS TO THE FIGURES

Figure 1. Panel A: Percent distribution of patients with or without accepted variants (AVs) according to the CH category. The percentage of positive cases tends to be lower in the patients with thyroid ectopy or athyreosis. Panel B: Percent distribution of AVs according to the type of CH and their association in single cases. Multiple gene involvements are seen in all CH categories.

Figure 2. Panel A: number of AVs with deleterious effect or prediction (loss-of-function by nature: frameshift, nonsense, splicing variants; and missense variants predicted to be deleterious by ≥5 out of 7 algorithms) identified in each candidate gene. Variants were indeed found more frequently in larger genes, such as *TG*, *DUOX2* or *TSHR*. Panel B: number of variants identified in each single gene according to the CH phenotype. Genes typically associated with functional defects (*TG, TPO, SLC26A4, TSHR, DUOX2*) were also found mutated in cases with thyroid dysgenesis, and viceversa genes typically associated with dysgenesis (*FOXE1, JAG1, NKX2-1*) were found mutated in cases with GIS.

Figure 3. Distribution of accepted variants (AVs) in 11 CH-related genes (*NKX2-1, PAX8, FOXE1, GLIS3, JAG1, TSHR, SLC26A4, TG, TPO, DUOX2, DUOXA2*) and 18 CH-unrelated genes (*AIP, CDKN1B, FGFR1, FOXL2, FSHR, GDF9, GH1, GHRHR, GNRH1, GNRH2, GNRHR, MEN1, NR5A1, PROK2, PROKR2, PROP1, POU1F1, RET*) among the CH patients and the control population. The CH patients are highly significantly enriched only with variations in CH-related genes. The significance is even stronger when variants with a demonstrated functional impact or predicted to be deleterious in ≥5/7 in silico programmes (disruptive AVs) are considered.

Figure 4. Distribution of accepted gene variants in nine families. The filled sections in the squares and circles outline the variants. All the CH probands (indicated by an arrow) were carriers of >1 variant (up to 5 variants in case 9 and his sister, both affected with CH) that occurred de novo or were inherited either from the mother or father. The probands 11, 12 and 20 had been previously reported (10, 35). The combinations of AVs in morphogenetic and functional genes represent a

suitable explanation for the thyroid dysgenesis seen in cases 11, 12 and 17. The variable combinations of multiple variants are associated with CH, whereas milder forms of non-autoimmune hypothyroidism were found in the relatives carrying only part of the variants present in the family (as in case 20, 33, 72 or 103). Familiarity for CH was known in cases 94 and 96 and both affected siblings carry the same combinations of AVs. The young sister of proband 103 (2 years old), the two brothers (12 and 2 years old) and both parents of proband 33, as well as the mother of case 72 were found to be hypothyroid after positive genetic screening. The co-segregation of CH with multiple gene defects supports the oligogenic model of CH. ASP: atrial septum defect; VSP: ventricular septum defect; TGV: transposition of great vessels.

Table 1. Clinical and biochemical characteristics of the 177 CH patients. Most of them had severe forms of CH at diagnosis. Thyroid diseases in the relatives were frequent and included nodular and autoimmune diseases in most of them. Instead familiarity for CH was reported in 6 CH cases, consistent with the prevalent sporadic presentation of the disease.

Characteristics	**N (%)**
Thyroid dysgenesis	**83** (46.9)
- athyreosis	23 (13.0)
- ectopy	30 (16.95)
- hypoplasia	30 (16.95)
Gland-in-situ (GIS)	**94** (53.1)
Severe CH (high TSH, low FT4)	**151** (85.3)
Mild CH (high TSH, low-normal FT4)	**26** (14.7)
Thyroid disease(s) in the family	**88** (49.7)
Associated abnormalities:	**55** (31.1)
- Nervous system	17 (9.6)
- Cardiovascular system	20 (11.3)
- Growth/skeletal	22 (12.4)
- Others	17 (9.6)

ABBREVIATIONS

Congenital Hypothyroidism: CH; Next Generation Sequencing: NGS; Whole Exome Sequencing: WES; Accepted Variant: AV; Gland-in-situ: GIS; Minor Allele Frequency: MAF; Loss-of-function: LOF; Free Thyroxine: FT4; Thyrotropin: TSH; Atrial Septum defect: ASD; Ventricular Septum Defect: VSD; Transposition of Great Vessels: TGV.

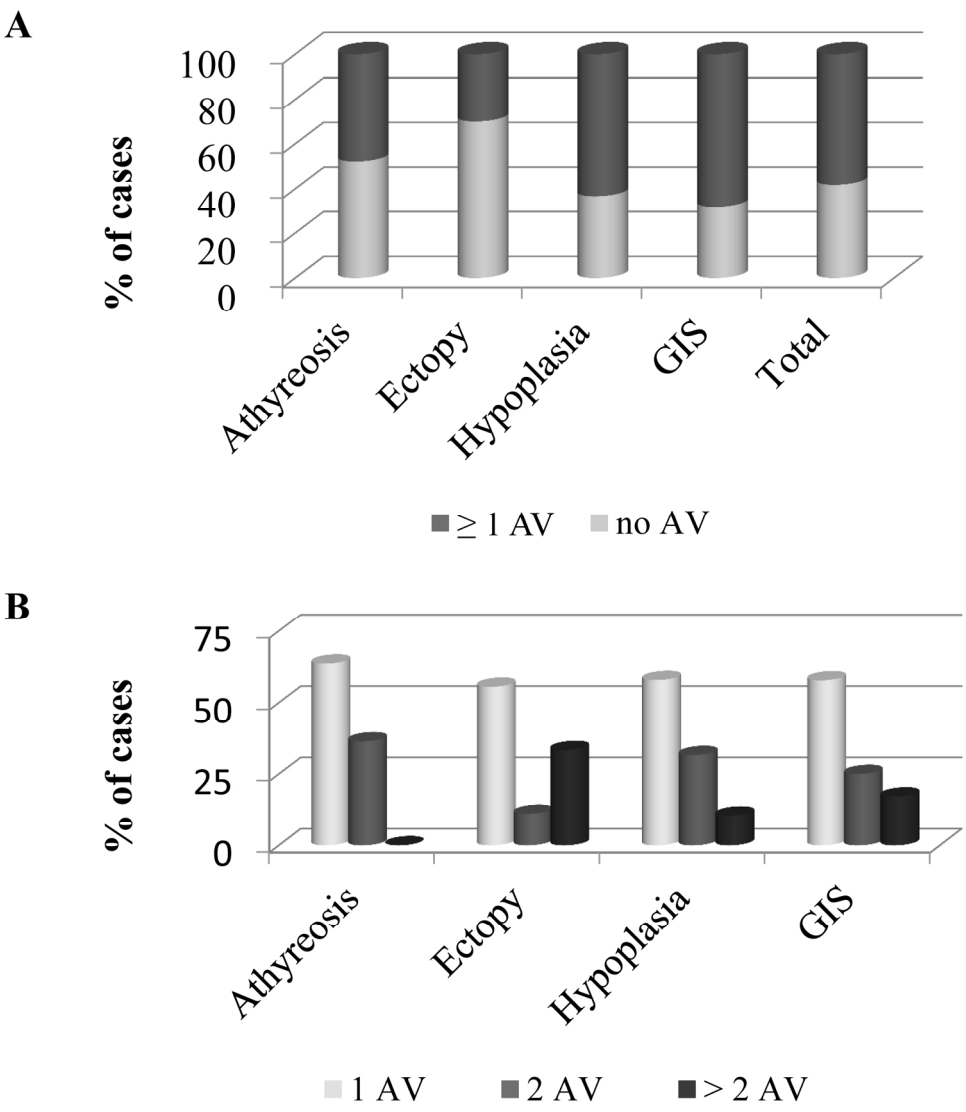

Figure 1. Panel A: Percent distribution of patients with or without accepted variants (AVs) according to the CH category. The percentage of positive cases tends to be lower in the patients with thyroid ectopy or athyreosis. Panel B: Percent distribution of AVs according to the type of CH and their association in single cases. Multiple gene involvements are seen in all CH categories.

162x182mm (300 x 300 DPI)

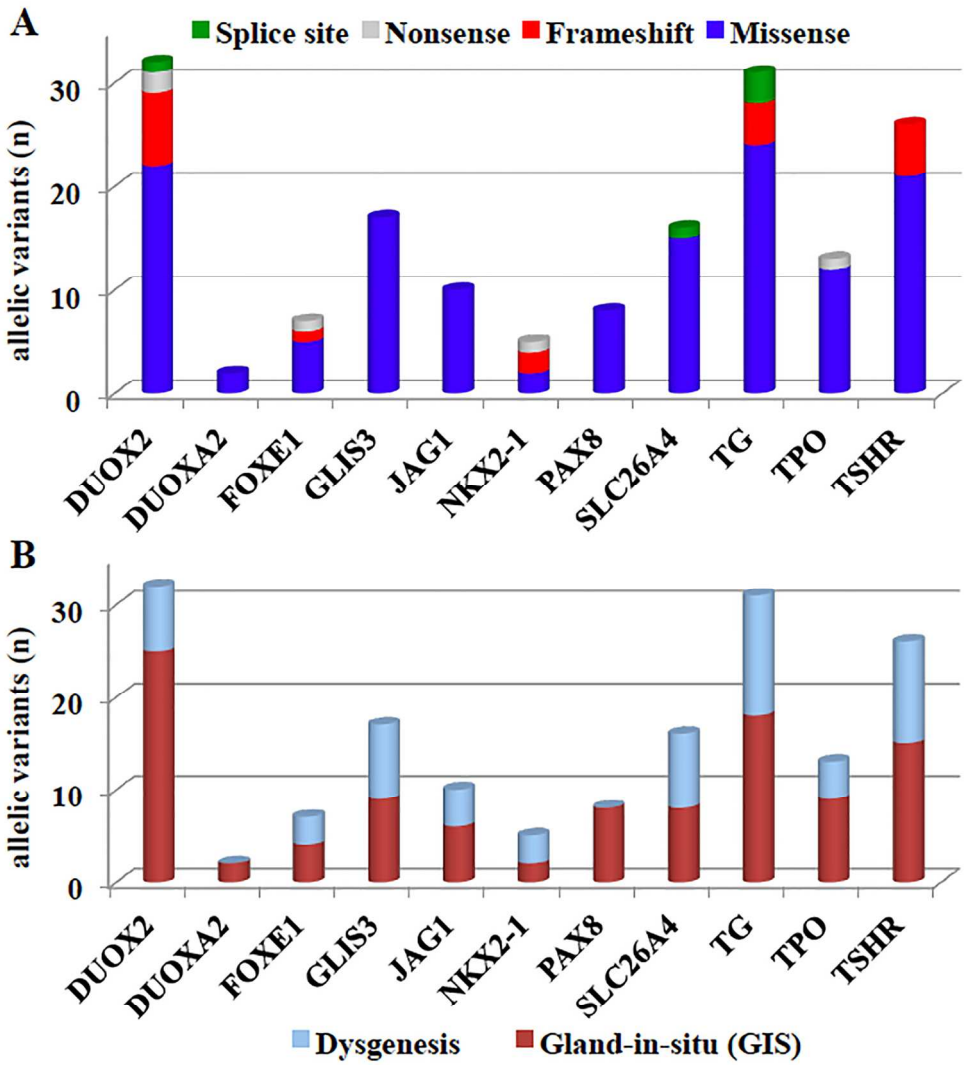

Figure 2. Panel A: number of AVs with deleterious effect or prediction (loss-of-function by nature: frameshift, nonsense, splicing variants; and missense variants predicted to be deleterious by ≥5 out of 7 algorithms) identified in each candidate gene. Variants were indeed found more frequently in larger genes, such as TG, DUOX2 or TSHR. Panel B: number of variants identified in each single gene according to the CH phenotype. Genes typically associated with functional defects (TG, TPO, SLC26A4, TSHR, DUOX2) were also found mutated in cases with thyroid dysgenesis, and viceversa genes typically associated with dysgenesis (FOXE1, JAG1, NKX2-1) were found mutated in cases with GIS.

170x194mm (300 x 300 DPI)

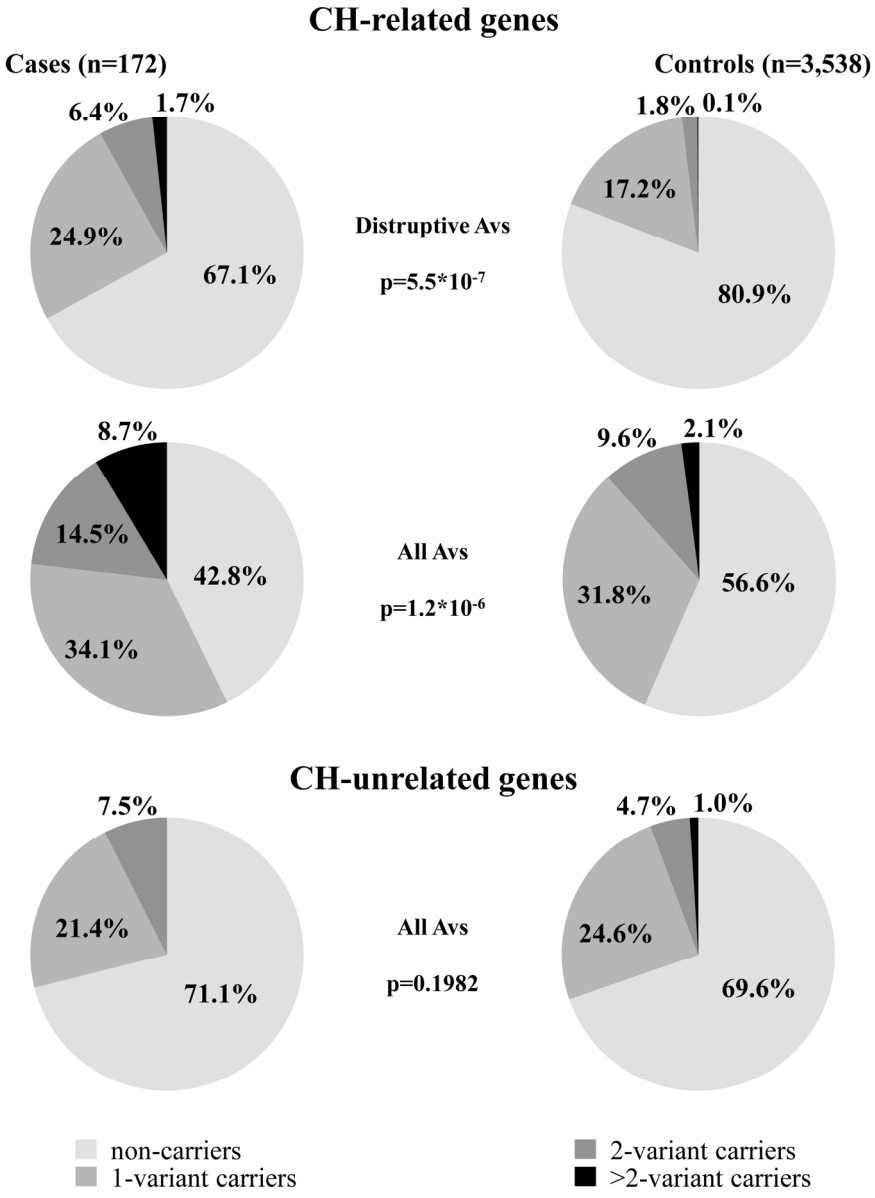

Figure 3. Distribution of accepted variants (AVs) in 11 CH-related genes (NKX2-1, PAX8, FOXE1, GLIS3, JAG1, TSHR, SLC26A4, TG, TPO, DUOX2, DUOXA2) and 18 CH-unrelated genes (AIP, CDKN1B, FGFR1, FOXL2, FSHR, GDF9, GH1, GHRHR, GNRH1, GNRH2, GNRHR, MEN1, NR5A1, PROK2, PROKR2, PROP1, POU1F1, RET) among the CH patients and the control population. The CH patients are highly significantly enriched only with variations in CH-related genes. The significance is even stronger when variants with a demonstrated functional impact or predicted to be deleterious in ≥5/7 in silico programmes (disruptive AVs) are considered.

173x234mm (300 x 300 DPI)

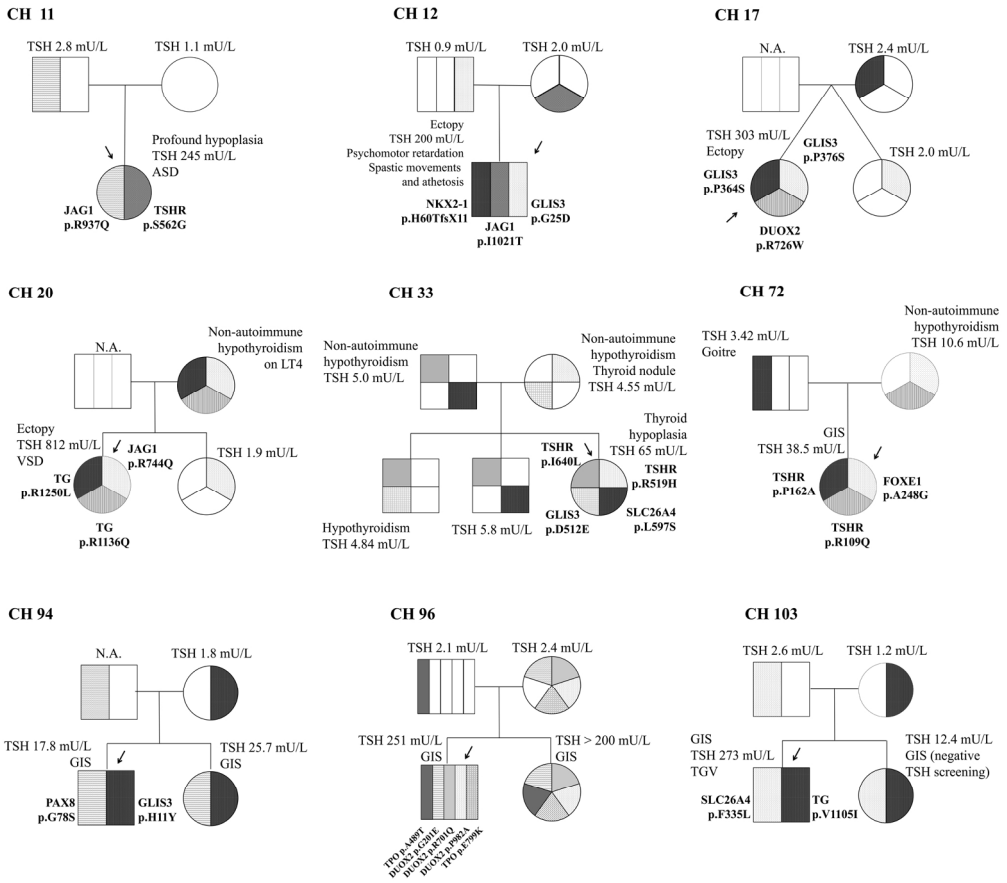

Figure 4. Distribution of accepted gene variants in nine families. The filled sections in the squares and circles outline the variants. All the CH probands (indicated by an arrow) were carriers of >1 variant (up to 5 variants in case 9 and his sister, both affected with CH) that occurred de novo or were inherited either from the mother or father. The probands 11, 12 and 20 had been previously reported (10, 35). The combinations of AVs in morphogenetic and functional genes represent a suitable explanation for the thyroid dysgenesis seen in cases 11, 12 and 17. The variable combinations of multiple variants are associated with CH, whereas milder forms of non-autoimmune hypothyroidism were found in the relatives carrying only part of the variants present in the family (as in case 20, 33, 72 or 103). Familiarity for CH was known in cases 94 and 96 and both affected siblings carry the same combinations of AVs. The young sister of proband 103 (2 years old), the two brothers (12 and 2 years old) and both parents of proband 33, as well as the mother of case 72 were found to be hypothyroid after positive genetic screening. The co-segregation of CH with multiple gene defects supports the oligogenic model of CH. ASD: atrial septum defect; VSD: ventricular septum defect; TGV: transposition of great vessels.

173x153mm (300 x 300 DPI)