Deciphering the mystery of CHNG3

Satoshi Narumi

Department of Pediatrics, Keio University School of Medicine, Tokyo, Japan

Congenital hypothyroidism (CH), characterized by insufficient thyroid hormone production due to abnormalities in the hypothalamic-pituitary-thyroid axis, is the most common congenital endocrine disorder. We previously conducted comprehensive genetic screening of 102 patients with permanent CH born in Kanagawa Prefecture, Japan and identified mutations in several genes in 19 CH patients, including defects in genes encoding dual oxidase 2, thyroglobulin, thyrotropin receptor, thyroid peroxidase, and paired-box 8. Despite these findings, approximately 80% of cases remain unexplained. CH pedigrees unexplained by known genetic forms of CH have been reported in the literature and registered as congenital hypothyroidism, nongoitrous, 3 (CHNG3; %609893) in Online Mendelian Inheritance in Man. We also identified a Japanese pedigree of CH that was compatible with CHNG3. However, the exact genetic cause of CHNG3 was not revealed by standard analysis methods such as exome sequencing and array comparative genomic hybridization. We therefore took a combined approach and analyzed a total of 11 undiagnosed CH pedigrees by whole genome sequencing to analyze a 3-Mb linkage region, and found a disease-causing variant affecting a TTTG microsatellite in a noncoding region on chromosome 15. Further analysis revealed that 13.9% of 989 Japanese CH patients had abnormalities involving the TTTG microsatellite, with a substantial proportion (41.5%) of familial CH cases carrying these mutations. Identification of the genetic cause of CHNG3 provides new insights into the pathogenesis of CH, and highlights the need for continued exploration of noncoding genomic regions in Mendelian disorders of unknown etiology.

Keywords: Congenital hypothyroidism, Genetics, Whole genome sequencing, Genetic linkage

Address for correspondence:
Satoshi Narumi
Department of Pediatrics, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan
Email: narumi-s@keio.jp
https://orcid.org/0000-0002-0940-4517

Highlights

· Previous studies have shown that approximately 20% of congenital hypothyroidism cases are due to mutations of single genes, and disease-causing variants have been detected in the coding region of these responsible genes.
· Using a combined linkage analysis and whole genome sequencing approach, a novel Mendelian congenital hypothyroidism (i.e., congenital hypothyroidism, nongoitrous, 3) due to variants in a noncoding region on chromosome 15 was discovered.
· These findings emphasize the need to explore noncoding genomic regions for Mendelian disorders, and highlight the effectiveness of combining whole genome sequencing with physical mapping methods.

Introduction

Disorders that result in deficient thyroid hormone production due to congenital abnormalities of the hypothalamus-pituitary-thyroid axis are collectively referred to as congenital hypothyroidism (CH). Primary CH is due to abnormalities of the thyroid gland,

while central CH is due to abnormalities in the hypothalamus or pituitary gland. Many countries, including Japan and Korea, screen newborns for CH based on measurement of thyroid-stimulating hormone (TSH) level on filter paper. However, only primary CH cases can be captured by TSH-based screening methods. In this review, unless otherwise noted, CH refers to primary CH.

There are 2 major clinical classifications of CH cases. One is based on the clinical course, distinguishing between permanent CH, where hypothyroidism persists throughout life, versus transient CH, where hypothyroidism is limited to early childhood. Nagasaki et al.[1] analyzed 240 patients diagnosed with CH through the newborn screening program who were treated with levothyroxine, and found that the ratio of patients with permanent to transient CH was approximately 3:1. Another clinical classification is based on thyroid morphology. In our studies of permanent CH patients, thyroid morphology was evaluated in 90 patients by thyroid ultrasonography or ^{123}I scintigraphy, and we found 41% had thyroid ectopy, 7% had thyroid aplasia, 9% had thyroid hypoplasia, 11% had goiter, and 32% had normal morphology.[2-4] Thyroid ectopy, aplasia and hypoplasia are collectively referred to as thyroid dysgenesis. Due to an absolute deficiency of thyroid hormone-producing cells, thyroid dysgenesis usually results in permanent CH. However, CH with goiter can also occur due to maternal anti-thyroid drug usage or excessive iodine exposure in the perinatal period. Thus, CH cases with goiter sometimes manifest as transient CH. Since information on thyroid morphology provides clinical clues to predict long-term outcomes, it is recommended that thyroid morphology be assessed by ultrasonography prior to initiating levothyroxine therapy.

Comprehensive genetic analysis of CH

CH is the most common congenital endocrine disorder, with a frequency of approximately 1 in 2,000 to 3,000 live births worldwide.[5] Some congenital endocrine disorders, such as congenital adrenal hyperplasia, are mostly due to Mendelian disorders, while others, such as congenital hypopituitarism, are mostly non-genetic diseases. What about CH?

We performed a comprehensive genetic screen of 102 patients with permanent CH drawn from the general population of 353,000 children in Kanagawa Prefecture, Japan, and demonstrated that 19 of the 102 subjects had genetic forms of CH.[2-4] Genetic causes of CH included mutations in the dual oxidase 2 (DUOX2) gene in 7 patients, thyroglobulin (TG) gene in 5 patients, thyrotropin receptor (TSHR) gene in 3 patients, thyroid peroxidase gene (TPO) in 2 patients, and paired-box 8 (PAX8) gene in 2 patients. Of the responsible genes identified in this cohort, only PAX8 defects showed autosomal dominant inheritance, while the other 4 genes showed autosomal recessive inheritance. These series of population-based genetic epidemiology studies revealed that approximately 20% of CH are known Mendelian CH cases, but the underlying cause remained unknown for approximately 80% of patients.

Discovery of CHNG3: European pedigrees

Table 1 shows a list of genes associated with primary CH that were identified as of 2024. As can be seen from this list, the major genes responsible for CH had been identified by 2008. These genes are transcribed into mRNA predominantly in the thyroid gland, and studies in cellular and animal models have elucidated the physiological roles that the gene products play in the thyroid (Fig. 1). It is generally believed that the turning point in human genetics research was 2009, when whole exome sequencing (WES) findings were first reported.[6] For example, identification of the immunoglobulin superfamily 1 (IGSF1) gene, the gene most frequently associated with central CH, was achieved by a study using exome sequencing in 2012.[7] In contrast, the only gene responsible for primary CH identified by WES is the solute carrier family 26, member 7 gene (SLC26A7).[8-10] Thus, research on the genetics of primary CH was regarded as "almost complete" before the introduction of WES. However, there was a report of 5 European CH families with autosomal dominant inheritance that could not be explained as known Mendelian forms by Grasberger et al. in 2005.[11,12] Analysis of the 5 families using genetic markers showed significant linkage to

Table 1. Responsible genes for Mendelian CH

Classification	Gene name*	Function of the gene product	Inheritance	Extrathyroidal manifestations	First reported
Goitrous CH	DUOX2	Membrane-bound enzyme	AR	None	2002
	TG	Hormone precursor	AR	None	1991
	TPO	Membrane-bound enzyme	AR	None	1992
	SLC5A5	Iodine transporter	AR	None	1997
	DUOXA2	Membrane-bound enzyme	AR	None	2008
	SLC26A4	Iodine transporter	AR	Sensorineural hearing loss (Pendred syndrome)	1997
	SLC26A7	Iodine transporter	AR	None	2018
	IYD	Membrane-bound enzyme	AR	None	2008
Nongoitrous CH	TSHR	Receptor	AR	None	1995
	PAX8	Transcription factor	AD	Congenital anomalies of the kidney and urinary tract (rare)	1998
	NKX2-1	Transcription factor	AD	Choreoathetosis, respiratory distress syndrome	1998
	FOXE1	Transcription factor	AR	Spiky hair, cleft palate	1998

CH, congenital hypothyroidism; AR, autosomal recessive; AD, autosomal dominant.
*The gene names are listed in order of frequency for goitrous CH and nongoitrous CH, respectively.

the long arm of chromosome 15. The disease was registered as congenital hypothyroidism, nongoitrous, 3 (CHNG3; %609893) in Online Mendelian Inheritance in Man (https://omim.org/), an online database of Mendelian genetic disorders. The clinical phenotype of CHNG3 is mild to moderate CH with normal
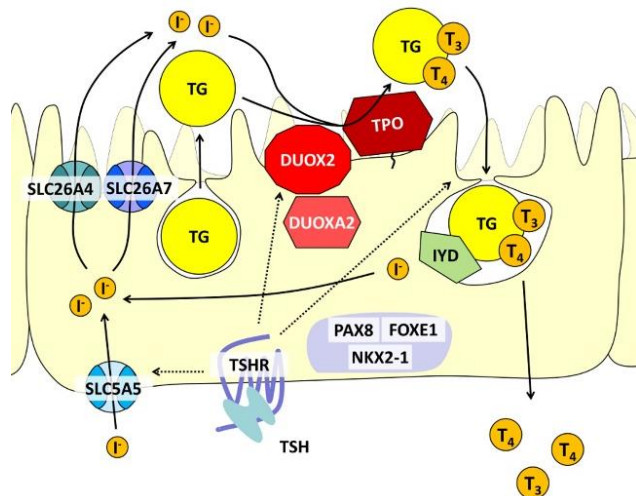

Fig. 1. Physiological roles of the causative genes of Mendelian congenital hypothyroidism in thyroid cells. One thyroid cell is drawn in the figure, with the vascular lumen at the bottom and the follicular lumen at the top. Inorganic iodine (I-) that reaches the thyroid gland via the blood stream is concentrated in thyroid cells by the action of sodium-iodine symporter (SLC5A5) and then pumped into the thyroid follicles by the action of SLC26A family transporters (SLC26A4 and SLC26A7). Thyroid follicles contain large amounts of thyroglobulin (TG), a macroprotein synthesized by thyroid cells. Thyroid peroxidase (TPO) catalyzes (i) the addition of inorganic iodine to several tyrosine residues of TG and (ii) a coupling reaction that creates a triiodothyronine (T3) or thyroxine (T4) group from 2 iodinated tyrosine residues. DUOXA2 is an endoplasmic reticulum protein that promotes the maturation of DUOX2. DUOX2 provides the hydrogen peroxide required for the iodination reaction of TG by TPO. Iodinated TG is taken up into thyroid cells by endocytosis and undergoes proteolysis in lysosomes to release T3 and T4. The iodotyrosine that is not converted to T3 and T4 is deiodinated by the membrane enzyme iodotyrosine deiodinase (IYD), and the recovered inorganic iodine is reused for thyroid hormone synthesis. These thyroid hormone synthesis pathways are positively regulated by the TSH-TSH receptor pathway. In addition, thyroid-specific transcription factors PAX8, NKX2-1, and FOXE1 regulate the expression levels of these thyroid hormone synthesis-related molecules. TSH, thyroid-stimulating hormone; TSHR, TSH receptor.

thyroid gland morphology.[11,12] Usually, once the disease-linked genomic region is specified, it does not take much time to identify the nucleotide-level causative DNA sequence change(s). In particular, when combined with WES, it is expected that the responsible gene will be identified efficiently. There are exceptional genomic abnormalities that cannot be detected by WES, such as exon-level deletions and variants in the regulatory regions or deep introns. However, it was considered unlikely that all 5 CHNG3 families reported by Grasberger et al.[12] had difficult-to-identify genomic abnormalities. However, even after WES became popular, these authors did not publish subsequent studies, suggesting some difficulties in follow-up.

CHNG3 in Japan

In 2006, we began a study of a Japanese CH pedigree showing autosomal dominant inheritance (Fig. 2).[13] Family members manifested with mild to moderate CH with normal thyroid morphology. The results of sequencing of known CH-related genes were negative. Using the GeneChip Mapping 250K Nsp Assay Kit (Affymetrix, Santa Clara, CA, USA), we found linkage to an approximately 3-Mb genomic region in chr15, which overlaps with the region reported by Grasberger et al.[12] This meant that a sixth CHNG3 pedigree was found in Japan. The linkage region encompassed 13 genes from the ATP/GTP-binding protein-like 1 gene (*AGBL1*) to the Fanconi anemia complementation group I (*FANCI*) in order from the centromere to the telomere. Since none of the genes were expressed predominantly in the thyroid, it was difficult to determine which was the most promising candidate. Follow-up analysis using array comparative genomic hybridization and WES were performed, but we could not identify the causative genomic abnormality. We also performed whole genome sequencing (WGS) assuming the causative genomic abnormality affected a regulatory region, such as an enhancer. As a result, more than 1,000 rare sequence variants were identified within the 3-Mb linkage region, including the true pathogenic variant that is present in the TTTG microsatellite, but it was impossible to determine which of these 1,000 candidate rare variants was the true cause.
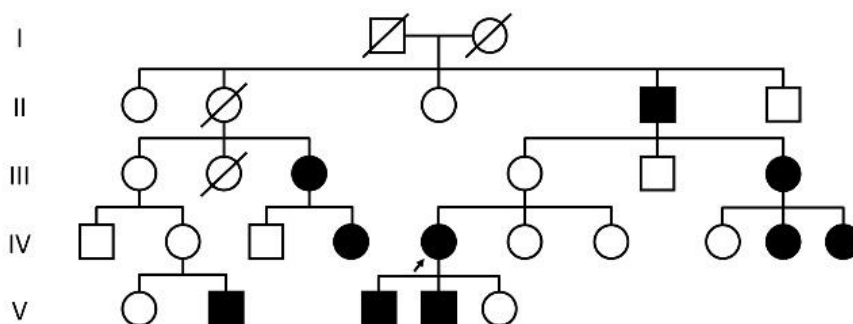

Fig. 2. A Japanese large pedigree of congenital hypothyroidism of unknown etiology. In the figure, square symbols indicate males and round symbols indicate females. Arrows indicate the proband. Filled symbols indicate individuals with hypothyroidism.

Subgroup analysis of Mendelian CH: a shift in approach

We started the molecular genetic study of CH in 2006, and took more than 15 years to accumulate data on about 1,000 Japanese CH patients. One day in 2022, through a subgroup analysis by presence or absence of family history of CH, we noticed that in the subgroup of patients with a positive family history, 22% had some form of genetic CH, which was not different from the proportion in the whole patient cohort (21%) (Fig. 3).[13] In clinical genetics, a positive family history is a key indicator of a Mendelian disorder. The fact that about 80% of CH cases with positive family history cannot be explained as known Mendelian CH suggested that there was still an unidentified Mendelian form of CH hidden, probably among the familial cases. If the cause of CHNG3 was a genomic abnormality difficult to detect by WES, then CHNG3 would have been among these undiagnosed familial CH cases.

To test this hypothesis, we analyzed 23 individuals from 10 undiagnosed CH families, in addition to 2 members of the largest family (Fig. 2) by WGS to search for a genomic region where common abnormalities were found. In typical human genetics research, only coding regions (genomic regions that specify the amino acid sequence of the gene product) are analyzed. In this research, the pathogenicity of identified sequence variants can be evaluated based on their presumed effect on the amino acid sequence, such as the introduction of a termination codon or change of an amino acid sequence. In contrast, analysis of noncoding regions does not allow for such an assessment of pathogenicity. For this reason, we simply compared the density of rare sequence variants between the patient group (N=25) and healthy controls (N=56), irrespective of the biological effects. The 3-Mb linkage region was divided into approximately 6,000 blocks of 500 bp each, and the density of rare variants, defined as an allele frequency of less than 1/500

in the 38KJPN database (https://jmorp.megabank.tohoku.ac.jp/), was compared between patients and controls. With this analysis, we aimed to identify blocks where rare variants were over-represented in undiagnosed familial CH patients. We identified only one block where rare variants were enriched in the patient group.[13] In the block, shortening of a TTTG microsatellite repeat (4 repeats in normal individuals and 3 in patients) was observed, which was shared by 8 of the 11 CH families analyzed. This microsatellite shortening was present in only 3 of 38,722 individuals in 38KJPN. In an analysis of 989 CH patients, 13.9% were found to have abnormalities involving the TTTG microsatellite (i.e., CHNG3). A subgroup analysis showed that as many as 41.5% of family history-positive CH patients had CHNG3. These results support a clinically plausible genetic explanation for the increased risk of Mendelian CH with a positive family history.

Conclusions

In 1995, Antonarakis et al.[14] estimated that 85% of all disease-causing genetic variants are found in coding regions. Then, in 2009, WES was introduced to human genetics researchers, and this new technology allowed the discovery of thousands of new Mendelian disease-related genomic abnormalities in coding regions. In the late 2010s, WGS, which can analyze noncoding regions as well, began to become more widely used, and as of 2024, it has become the standard genome analysis method along with WES. However, few Mendelian disorders that are primarily caused by abnormalities in noncoding regions have been discovered.

There are 2 main challenges. First, noncoding regions are very large, approximately 60 times larger than coding regions. The larger the region to be explored, the more difficult it is to identify the cause of the disease. A second challenge is the difficulty in inferring the biological effects of sequence variations in
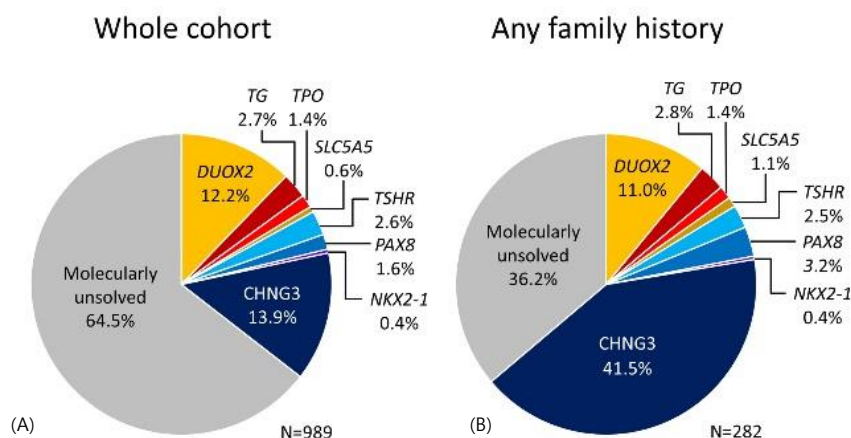

Fig. 3. Percentage of Mendelian forms of congenital hypothyroidism (CH). Prior to the recognition of CHNG3 (shown in dark blue in the figure), there was no difference in the proportion of patients with Mendelian forms of CH, which was about 20% in the whole patient cohort (A) and in the subgroup with any family history (B). Today, about 1/3 of all patients and about 2/3 of patients with family history have known Mendelian forms of CH. TG, thyroglobulin; TPO, thyroid peroxidase; CHNG3, congenital hypothyroidism, nongoitrous, 3. Adapted from Narumi et al. Nat Genet 2024;56:869-876.[13]

noncoding regions. The first challenge can be addressed with a 2-step strategy, as was done in CHNG3: (1) defining the candidate region using physical mapping methods such as linkage analysis; and (2) performing WGS of undiagnosed cases (ideally, familial cases) to identify the region(s) where rare sequence variants are over-represented in patients relative to controls. Physical mapping is a valuable method of identifying genes responsible for autosomal dominant traits, but since the introduction of WES, its value has decreased dramatically. However, now that the identification of responsible genes with WES has clearly peaked, it is time for physical mapping methods to be reevaluated. As for the second challenge, inference of the biological effects of sequence changes in noncoding regions has been the subject of intense research by genome researchers in relation to interpretation of signals from genome-wide association studies. Studies on epigenomic interactions between DNA, histone proteins and transcriptional regulators are in progress. These studies will greatly advance our understanding of the biological roles of noncoding regions.

Could it be that about 15% of Mendelian disorders are due to abnormalities in noncoding regions, as Cooper and colleagues have envisioned? We are optimistic that we will soon be able to answer this question.

Notes

ORCID

Satoshi Narumi: 0000-0002-0940-4517

References

1. Nagasaki K, Asami T, Ogawa Y, Kikuchi T, Uchiyama M. A study of the etiology of congenital hypothyroidism in the Niigata prefecture of Japan in patients born between 1989 and 2005 and evaluated at ages 5-19. Thyroid 2011;21:361-5.

2. Narumi S, Muroya K, Abe Y, Yasui M, Asakura Y, Adachi M, et al. TSHR mutations as a cause of congenital hypothyroidism in Japan: a population-based genetic epidemiology study. J Clin Endocrinol Metab 2009;94:1317-23.

3. Narumi S, Muroya K, Asakura Y, Aachi M, Hasegawa T. Molecular basis of thyroid dyshormonogenesis: genetic screening in population-based Japanese patients. J Clin Endocrinol Metab 2011,96:E1838-42.

4. Narumi S, Muroya K, Asakura Y, Adachi M, Hasegawa T.

Transcription factor mutations and congenital hypothyroidism: systematic genetic screening of a population-based cohort of Japanese patients. J Clin Endocrinol Metab 2010,95:1981-5.

5. Corbetta C, Weber G, Cortinovis F, Calebiro D, Passoni A, Vigone MC, et al. A 7-year experience with low blood TSH cutoff levels for neonatal screening reveals an unsuspected frequency of congenital hypothyroidism (CH). Clin Endocrinol (Oxf) 2009;71:739-45.

6. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 2009;461:272-6.

7. Sun Y, Bak B, Schoenmakers N, van Trotsenburg AS, Oostdijk W, Voshol P, et al. Loss-of-function mutations in IGSF1 cause an X-linked syndrome of central hypothyroidism and testicular enlargement. Nat Genet 2012;44:1375-81.

8. Zou M, Alzahrani AS, Al-Odaib A, Alqahtani MA, Babiker O, Al-Rijjal RA, et al. Molecular analysis of congenital hypothyroidism in Saudi Arabia: SLC26A7 mutation is a novel defect in thyroid dyshormonogenesis. J Clin Endocrinol Metab 2018;103:1889-98.

9. Cangul H, Liao XH, Schoenmakers E, Kero J, Barone S, Srichomkwun P, et al. Homozygous loss-of-function mutations in SLC26A7 cause goitrous congenital hypothyroidism. JCI Insight 2018;3:e99631.

10. Ishii J, Suzuki A, Kimura T, Tateyama M, Tanaka T, Yazawa T, et al. Congenital goitrous hypothyroidism is caused by dysfunction of the iodide transporter SLC26A7. Commun Biol 2019;2:270.

11. Grasberger H, Mimouni-Bloch A, Vantyghem MC, van Vliet G, Abramowicz M, Metzger DL, et al. Autosomal dominant resistance to thyrotropin as a distinct entity in five multigenerational kindreds: clinical characterization and exclusion of candidate loci. J Clin Endocrinol Metab 2005;90:4025-34.

12. Grasberger H, Vaxillaire M, Pannain S, Beck JC, Mimouni-Bloch A, Vatin V, et al. Identification of a locus for nongoitrous congenital hypothyroidism on chromosome 15q25.3-26.1. Hum Genet 2005;118:348-55.

13. Narumi S, Nagasaki K, Kiriya M, Uehara E, Akiba K, Tanase-Nakao K, et al. Functional variants in a TTTG microsatellite on 15q26.1 cause familial nonautoimmune thyroid abnormalities. Nat Genet 2024;56:869-76.

14. Antonarakis SE, Krawczak M, Cooper DN. The nature and mechanisms of human gene mutation. In: Sriver CR, Beaudet AL, Sly WS, Valle D, editors. The metabolic and molecular bases of inherited disease. 7th Ed. New York: McGraw-Hill; 1995. p. 259-91.