

Ester Saraiva Brust

**IDENTIFICAÇÃO DE ALTERAÇÕES GENÉTICAS NO EXOMA
COMPLETO DE PACIENTES COM HIPOTIREOIDISMO CONGÊNITO
POR DISGENESIA TIREOIDIANA**

Tese apresentada à Universidade Federal de São Paulo - Programa de Pós-graduação em Biotecnologia Interunidades, para obtenção do título de Doutora em Biotecnologia.

São Paulo
2022

Ester Saraiva Brust

**IDENTIFICAÇÃO DE ALTERAÇÕES GENÉTICAS NO EXOMA
COMPLETO DE PACIENTES COM HIPOTIREOIDISMO CONGÊNITO
POR DISGENESIA TIREOIDIANA**

Tese apresentada à Universidade Federal de São Paulo - Programa de Pós-graduação em Biotecnologia Interunidades, para obtenção do título de Doutora em Biotecnologia.

Orientadora: Prof^a Dr^a Ileana Gabriela Sánchez de Rubió

São Paulo

2022

Ficha catalográfica elaborada pela Biblioteca Prof. Antonio Rubino de Azevedo,
Campus São Paulo da Universidade Federal de São Paulo, com os dados fornecidos pelo(a) autor(a)

Saraiva Brust, Ester

IDENTIFICAÇÃO DE ALTERAÇÕES GENÉTICAS NO EXOMA
COMPLETO DE PACIENTES COM HIPOTIREOIDISMO CONGÊNITO
POR DISGENESIA TIREOIDIANA / Ester Saraiva Brust. - São Paulo,
2022.

xvi, 91f.

Tese (Doutorado) - Universidade Federal de São Paulo, Escola Paulista
de Medicina. Programa de Pós-Graduação em Biotecnologia Interunidades.

Título em inglês: Identification of genetic alterations in the complete
exome of patients with congenital hypothyroidism due to thyroid
dysgenesis.

1. Hipotireoidismo congênito. 2. Disgenesia tireoidiana. 3.
Sequenciamento do exoma. 4. Variantes.

UNIVERSIDADE FEDERAL DE SÃO PAULO
DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA
INTERUNIDADES

Chefe de departamento:

Prof Dr Marcelo Cristino Gama

Coordenadora do curso de pós-graduação:

Profª Drª Claudia Barbosa Ladeira Campos



Dedico este trabalho à minha filha Lara (Florzinha), que dividiu o tempo da mamãe com o doutorado. Foi intenso e trabalhoso, mas valeu todo o esforço para criar vocês duas (a Lara e a tese).

Agradecimentos

"Pois todas as coisas vêm dEle, existem por meio dEle e são para Ele. A Ele seja toda a glória para sempre! Amém". Romanos 11:36

Meu primeiro e maior agradecimento é à Deus, fonte de todo o conhecimento, força e persistência para o desenvolvimento deste trabalho, pois "quando sou fraco, então é que sou forte".

Agradeço a todos que estiveram perto (ou intimamente longe) durante todos estes anos de pesquisa. Familiares, amigos e irmãos sempre atenciosos e com os quais posso compartilhar essa conquista. Tantas coisas mudaram desde quando finalizei meu mestrado, tantas pessoas se foram e outras se achegaram, e agradeço a todos pois moldaram o que sou hoje.

À minha orientadora Dra Ileana Rubió, sempre esquecida, simples, exigente e humana. Sem sua direção não teríamos um trabalho tão rico e bem desenvolvido. Sua amizade e orientação começaram durante o meu mestrado e se estenderam até aqui (e espero que continuem, não importa para onde a vida nos levar a partir de agora).

A todos os colegas do Laboratório de Ciências Moleculares da Tireoide e Laboratório de Biologia Molecular do Câncer que me acompanharam e cooperaram com seu conhecimento e amizade. Em especial, às minhas companheiras Dra Erika Urbano e Dra Mariana Rodrigues e todos os outros alunos da prof^a Ileana. Ao Dr André Vettore e seus alunos que participaram no início do estudo e ao Dr Rodrigo Tamura e seus alunos que chegaram mais tarde. Todos foram importantes para a minha formação acadêmica.

Aos nossos colaboradores Dr Helton Ramos do Laboratório de Estudos da Tireoide (Universidade Federal da Bahia) e Dra Tatiana Amorim da APAE Salvador; Dr Geraldo Medeiros Neto (*in memoriam*) do Laboratório de Endocrinologia Celular e Molecular, Unidade de Tireoide LIM-25 (Faculdade de Medicina da Universidade de São Paulo); Dr Celso Friguglietti da Cirurgia de Cabeça e Pescoço (Hospital Santa Catarina de São Paulo); Dra Célia Nogueira do Departamento de Clínica Interna (Faculdade de Medicina de Botucatu, Universidade Estadual Paulista); Dr José Sgarbi do Departamento de Medicina, Divisão de Endocrinologia e Metabolismo (Faculdade de Medicina de Marília); Dra Anamaria Camargo, Dr Pedro Galante e Dr Filipe Ferreira

do Centro de Oncologia Molecular (Instituto de Ensino e Pesquisa do Hospital Sírio-Libanês); Dr Diego Mazzotti do Departamento de Psicobiologia da UNIFESP, São Paulo.

Este projeto foi financiado através do auxílio pesquisa FAPESP 2014/24549-4, e bolsas CAPES 88882.431003/2019-01 e CNPq (PIBIC 2018/2019).

Resumo

O hipotireoidismo congênito (HC) é um distúrbio metabólico sistêmico, onde a produção de T3 e T4 no período neonatal é insuficiente. Os pacientes com HC apresentam disormonogênese (defeitos de síntese ou secreção dos hormônios tireoidianos, afetando 10-15% dos casos) ou disgenesia tireoidiana (DT) (defeitos no desenvolvimento da tireoide, afetando 80-85% dos casos). A DT já foi relacionada a mutações nos genes *HOXA3*, *NKX2-1*, *PAX8*, *FOXE1*, *HHEX*, *HES1*, *EYA1*, *TSHR*, *TUBB1*, *GLIS3*, *JAG1*, *NKX2-5*, *NTN1* e *CDCA8*, embora tenham sido identificadas em apenas 5% dos pacientes. Por isso, é necessário continuar investigando as bases moleculares da embriogênese tireoidiana.

Objetivo: Ampliar o estudo das bases moleculares da DT através da pesquisa de alterações gênicas com relevância clínica no exoma completo e em sequências regulatórias, utilizando sequenciamento de nova geração e análises *in silico*.

Métodos: A casuística deste estudo foi composta por 71 pacientes com DT, 35 destes foram submetidos ao sequenciamento do exoma. As análises foram realizadas considerando genes alterados e variantes em comum nas DTs, variantes em genes candidatos (associados ao desenvolvimento e função tireoidiana) e regulados pelos fatores de transcrição FOXE1 e PAX8, e alterações cromossômicas. Informações das variantes sobre frequências populacionais e associação com doenças foram obtidas dos bancos de dados: 1000G, ABraOM, dbSNP, ExAc, ClinVar e DGV. Para a predição do efeito das variantes foram utilizados os programas: *Enrichr*, *FannsDB-Condol*, *FATHMM-MKL*, *HOPE*, *MutationAssessor*, *MutationTaster*, *PolyPhen2*, *String* e *TFBind*. A confirmação de variantes de interesse e análises de genótipo em pacientes e indivíduos controles normais brasileiros foram realizados por sequenciamento Sanger.

Resultados: Observamos em todos os pacientes genes alterados associados ao desenvolvimento, morfogênese, migração e adesão celular, assim como variantes raras (MAF<1%) e deletérias compartilhadas entre as DTs. Identificamos variantes raras e deletérias em genes candidatos, das quais se destacaram TSHR_p.T574S, FGF3_p.K101N e SFRP2_p.S235X. Também identificamos 5 variantes no gene *DUOX2* potencialmente prejudiciais e variantes em genes regulados por FOXE1 e PAX8, contudo, sem comprometer os *motifs* de ligação dos fatores de transcrição.

Nas análises de genotipagem verificamos que as variantes rs1867277 e rs71369530 no gene *FOXE1* foram fatores de risco para a DT (OR=2,45 (95%CI=1,69-3,55) e OR=1,78 (95%CI=1,10-2,88), respectivamente), com uma associação mais forte na presença conjunta das duas variantes (OR=2,90; 95%CI=1,24-6,72). Não identificamos CNVs ou regiões em homozigose prejudiciais. As análises de interação dos genes alterados sugerem uma possível causa poligênica da DT em nossos casos.

Conclusões: Este estudo permitiu amplamente rastrear variantes em casos de DT e identificar alterações com potencial relevância clínica. Ainda, polimorfismos no gene *FOXE1* foram fatores de risco para o desenvolvimento de malformações da tireoide. Nossos resultados sugerem que genes até então não associados às DTs podem estar envolvidos na doença e, ainda, reforçam a inexistência de um padrão genético único para as causas da DT. Os modelos poligênico ou de múltiplos fatores também podem estar relacionados com as falhas de desenvolvimento da tireoide, os quais devem ser mais explorados. Não descartamos a necessidade de ensaios funcionais e até mesmo o uso de modelos animais que podem fornecer dados mais conclusivos a associação destas variantes com a DT.

Abstract

Congenital hypothyroidism (CH) is a systemic metabolic disorder where the production of T3 and T4 in the neonatal period is insufficient. Patients with CH have dysmorphogenesis (defects in the synthesis or secretion of thyroid hormones, affecting 10-15% of cases) or thyroid dysgenesis (TD) (defects in thyroid development, affecting 80-85% of cases). TD has been linked to mutations in *HOXA3*, *NKX2-1*, *PAX8*, *FOXE1*, *HHEX*, *HES1*, *EYA1*, *TSHR*, *TUBB1*, *GLIS3*, *JAG1*, *NKX2-5*, *NTN1* and *CDCA8* genes, although they have been identified in only 5% of patients. Therefore, it is necessary to continue investigating the molecular basis of thyroid embryogenesis.

Objective: To expand the molecular basis study of TD through the investigation of genetic alterations with clinical relevance in the whole exome and in regulatory sequences, using next-generation sequencing and *in silico* analysis.

Methods: The sample of this study consisted of 71 patients with TD, 35 of whom underwent exome sequencing. The analyzes were performed considering altered genes and variants in common in TDs, variants in candidate genes (associated with thyroid development and function) and regulated by FOXE1 and PAX8 transcription factors, and chromosomal alterations. Variant information on population frequencies and disease association was obtained from the following databases: 1000G, ABraOM, dbSNP, ExAc, ClinVar and DGV. To predict the variants effect, the following programs were used: Enrichr, FannsDB-Condol, FATHMM-MKL, HOPE, MutationAssessor, MutationTaster, PolyPhen2, String and TFBind. Confirmation of variants of interest and genotype analyzes in patients and normal Brazilian controls subjects were performed by Sanger sequencing.

Results: We observed in all patients altered genes associated with development, morphogenesis, migration and cell adhesion, as well as rare (MAF<1%) and deleterious variants shared between TDs. We identified rare and deleterious variants in candidate genes, of which TSHR_p.T574S, FGF3_p.K101N and SFRP2_p.S235X stood out. We also identified 5 potentially harmful variants in the *DUOX2* gene and variants in genes regulated by FOXE1 and PAX8, however, without compromising the transcription factor binding motifs. In the genotyping analysis, we found that the rs1867277 and rs71369530 variants in the *FOXE1* gene were risk factors for TD

(OR=2.45 (95%CI=1.69-3.55) and OR=1.78 (95%CI =1.10-2.88), respectively), with a stronger association in the joint presence of the two variants (OR=2.90; 95%CI=1.24-6.72). We did not identify harmful CNVs or homozygous regions. The interaction analyzes of the altered genes suggest a possible polygenic cause of TD in our cases. **Conclusions:** This study made it possible to broadly screen variants in TD cases and identify changes with potential clinical relevance. Furthermore, polymorphisms in the *FOXE1* gene were risk factors for the development of thyroid malformations. Our results suggest that genes not previously associated with TDs may be involved in the disease and, furthermore, reinforce the inexistence of a single genetic pattern for the causes of TD. Polygenic or multiple factor models may also be related to thyroid development failures, which should be further explored. We do not rule out the need for functional assays and even the use of animal models that can provide more conclusive data on the association of these variants with TD.

Sumário

Lista de figuras.....	xiii
Lista de tabelas.....	xv
1 INTRODUÇÃO.....	1
1.1 Tireoide.....	1
1.1.1 Desenvolvimento da tireoide.....	1
1.2 Hipotireoidismo congênito.....	3
1.2.1 Disgenesia tireoidiana.....	4
1.3 Sequenciamento de nova geração.....	7
2 OBJETIVO.....	9
2.1 Objetivos específicos.....	9
3 MÉTODOS.....	10
3.1 Pacientes.....	10
3.2 Sequenciamento.....	10
3.2.1 Sequenciamento do exoma.....	11
3.2.2 Sequenciamento Sanger.....	12
3.3 Análise dos dados de sequenciamento do exoma.....	12
3.3.1 Análise de variantes.....	12
3.3.2 Análise de CNVs.....	13
3.4 Ferramentas de bioanálise.....	13
3.4.1 Bancos de dados.....	14
3.4.2 Programas de bioinformática.....	14
3.4.3 Análises estatísticas.....	15
4 RESULTADOS.....	16
4.1 Causas genéticas em comum.....	16
4.1.1 Genes alterados em comum.....	16
4.1.2 Variantes em comum.....	19
4.2 Variantes em genes candidatos.....	23
4.2.1 Análises <i>in silico</i> de variantes em genes candidatos.....	27
4.2.2 Genes regulados por fatores de transcrição.....	33
4.2.3 Análise do gene <i>FOXE1</i>	37
4.2.4 Análise da família DUOX.....	44

4.3 Regiões em homozigose	57
4.4 Variantes estruturais CNVs	59
4.5 Análise individual dos pacientes	60
5 DISCUSSÃO	64
6 CONCLUSÕES	76
7 REFERÊNCIAS.....	77
Anexos	92

Lista de figuras

Figura 1: Fases do desenvolvimento embrionário da tireoide humana e de camundongo.....	2
Figura 2: Diagrama de Venn mostrando o número de genes alterados no exoma das 35 amostras de DT sequenciadas.....	17
Figura 3: Diagrama de Venn mostrando o número de variantes em comum no exoma das 35 amostras de DT sequenciadas.....	19
Figura 4: Diagrama de Venn mostrando as 31 variantes raras e deletérias em comum no exoma das 35 amostras de DT sequenciadas.....	20
Figura 5: Interação dos 7 genes com variantes raras e deletérias em comum identificadas no exoma das 35 amostras de DT sequenciadas.....	23
Figura 6: Interação dos genes candidatos alterados (análise por paciente) gerada pela ferramenta <i>String</i>	27
Figura 7: Quadro informativo com dados gerados pela ferramenta <i>MutationTaster</i> sobre a variante p.T574S no gene <i>TSHR</i>	28
Figura 8: Imagens geradas pela ferramenta <i>HOPE</i> para a troca <i>TSHR_p.T574S</i>	29
Figura 9: Quadro informativo com dados gerados pela ferramenta <i>MutationTaster</i> sobre a variante p.K101N no gene <i>FGF3</i>	30
Figura 10: Imagens geradas pela ferramenta <i>HOPE</i> para a troca <i>FGF3_p.K101N</i> ..	31
Figura 11: Quadro informativo com dados gerados pela ferramenta <i>MutationTaster</i> sobre a variante p.S235X no gene <i>SFRP2</i>	32
Figura 12: Representação esquemática da proteína DUOX2 na membrana apical do tireócito com a localização aproximada das variantes identificadas e confirmadas nas disgenesias tireoidianas.	46
Figura 13: Quadro informativo com dados gerados pela ferramenta <i>MutationTaster</i> sobre a variante p.T423I no gene <i>DUOX2</i>	47
Figura 14: Representação molecular dos aminoácidos treonina e isoleucina na troca <i>DUOX2_p.T423I</i>	49
Figura 15: Quadro informativo com dados gerados pela ferramenta <i>MutationTaster</i> sobre a variante p.R516H no gene <i>DUOX2</i>	49
Figura 16: Representação molecular dos aminoácidos arginina e histidina na troca <i>DUOX2_p. p.R516H</i>	50

Figura 17: Quadro informativo com dados gerados pela ferramenta <i>MutaionTaster</i> sobre a variante p.F966SfsX29 no gene <i>DUOX2</i>	52
Figura 18: Quadro informativo com dados gerados pela ferramenta <i>MutaionTaster</i> sobre a variante p.A1123T no gene <i>DUOX2</i>	53
Figura 19: Representação molecular dos aminoácidos alanina e treonina na troca <i>DUOX2_p. A1123T</i>	54
Figura 20: Quadro informativo com dados gerados pela ferramenta <i>MutaionTaster</i> sobre a variante p.N1170T no gene <i>DUOX2</i>	56
Figura 21: Representação molecular dos aminoácidos asparagina e treonina na troca <i>DUOX2_p.N1170T</i>	57
Figura 22: Distribuição das principais variantes identificadas no exoma das 35 amostras de DT sequenciadas	61
Figura 23: Interação dos genes alterados (análise por paciente) gerada pela ferramenta <i>String</i> . Pacientes com agenesia.....	62
Figura 24: Interação dos genes alterados (análise por paciente) gerada pela ferramenta <i>String</i> . Pacientes com hipoplasia62	
Figura 25: Interação dos genes alterados (análise por paciente) gerada pela ferramenta <i>String</i> . Pacientes com ectopia.....	63

Lista de tabelas

Tabela 1: Termos selecionados resultantes da pesquisa feita dentro da ferramenta <i>Enrichr</i> dos 2.393 genes alterados em comum no exoma das 35 amostras de DT sequenciadas (termos com $p < 0,05$) (continua)	17
Tabela 2: Variantes raras e deletórias em comum identificadas no exoma das 35 amostras de DT sequenciadas (continua).....	20
Tabela 3: Termos selecionados resultantes da pesquisa feita dentro da ferramenta <i>Enrichr</i> dos 7 genes com variantes raras e deletórias em comum identificadas no exoma das 35 amostras de DT sequenciadas (termos com $p < 0,05$)	22
Tabela 4: Variantes raras e deletórias em genes candidatos identificadas no exoma das 35 amostras de DT sequenciadas (continua).....	24
Tabela 5: Variantes raras na região 5'UTR de genes com motif para os fatores de transcrição FOXE1 e PAX8 identificadas nas amostras de disgenesias sequenciadas por WES	34
Tabela 6: Variantes raras e deletórias na região codificadora de genes com motif para os fatores de transcrição FOXE1 e PAX8 identificadas no exoma das 35 amostras de DT sequenciadas (continua).....	35
Tabela 7: Variantes identificadas na região de FOXE1 no exoma das 35 amostras de DT sequenciadas (continua).....	37
Tabela 8: Genótipos identificados nas disgenesias, controles normais e populações das bases de dados <i>ABraOM</i> e <i>1000G</i> para os polimorfismos rs907577, rs1867277 e rs71369530 do gene <i>FOXE1</i>	39
Tabela 9: Análise dos polimorfismos rs907577, rs1867277 e rs71369530 do gene <i>FOXE1</i>	40
Tabela 10: Análise conjunta dos polimorfismos rs1867277 e rs71369530 do gene <i>FOXE1</i>	42
Tabela 11: Análise da frequência dos polimorfismos rs7849497, rs1867277, rs1867278, rs1867279, rs1867280 e rs71369530 do gene <i>FOXE1</i> no exoma das 35 amostras de DT sequenciadas	43
Tabela 12: Variantes na região 3'UTR de FOXE1 no exoma das 35 disgenesias sequenciadas presentes em sítios de ligação de miRNAs (continua).....	43

Tabela 13: Variantes deletérias selecionadas na região dos genes da família DUOX identificadas no exoma das 35 amostras de DT sequenciadas	45
Tabela 14: Variantes contidas nas regiões em homozigose identificadas no exoma das amostras sequenciadas dos pacientes 14 e 24 sequenciados por WES	58
Tabela 15: Alterações estruturais (CNVs) identificadas no exoma das amostras de DT sequenciadas.....	59

1 INTRODUÇÃO

1.1 Tireoide

A principal função da glândula tireoide é a produção dos hormônios triiodotironina (T3) e tetraiodotironina ou tiroxina (T4). Estes hormônios promovem a regulação do consumo energético no organismo, sendo indispensáveis para o crescimento, o desenvolvimento e a maturação de vários órgãos (1).

Nos mamíferos, a glândula está localizada na região anterior do pescoço e possui dois tipos celulares distintos: células foliculares, que produzem os hormônios T3 e T4, e células parafoliculares ou células C, produtoras de calcitonina. As células foliculares são as mais numerosas e agrupam-se em estruturas esféricas de única camada celular chamadas folículos, onde são produzidos e armazenados os hormônios tireoidianos (HT). As células C ficam dispostas entre os folículos em posição parafolicular (2).

1.1.1 Desenvolvimento da tireoide

A glândula tireoide é a primeira estrutura endócrina que se desenvolve no embrião humano, a organogênese da glândula inicia-se após 22 dias da concepção e finaliza no 51º dia. A tireoide é bilobular e formada pelos dois tipos celulares distintos. As células foliculares originam-se a partir de um conjunto restrito de células do epitélio endodérmico do assoalho da faringe primitiva localizado no forâmen cécum, na base da língua em desenvolvimento. Essas células formam um espessamento que prolifera e evagina formando o divertículo tireoidiano, que invade o mesênquima ao seu redor e migra caudalmente em direção ventrodorsal e anteroposterior através da linha média anterior do pescoço. Durante a migração a tireoide adquire estrutura bilobular e o duto tireoglosso, porção de tecido que conectava a tireoide ao sítio de origem, desaparece. Quando a tireoide alcança sua posição final na frente da traqueia na base do pescoço, ocorre a incorporação dos corpos ectodermis ultimobrânquiais, dos quais se diferenciam as células C. A porção mediana da tireoide em formação permanece como istmo conectando os dois lobos através da linha média. O passo final da morfogênese é a organização dos folículos ou foliculogênese. Ao longo do

desenvolvimento da tireoide ocorre a ativação do programa de diferenciação funcional, que promove a expressão de genes específicos da tireoide (*TSHR*, *SLC5A5*, *TPO* e *TG*) que permitem a síntese hormonal (2-6) (Figura 1).

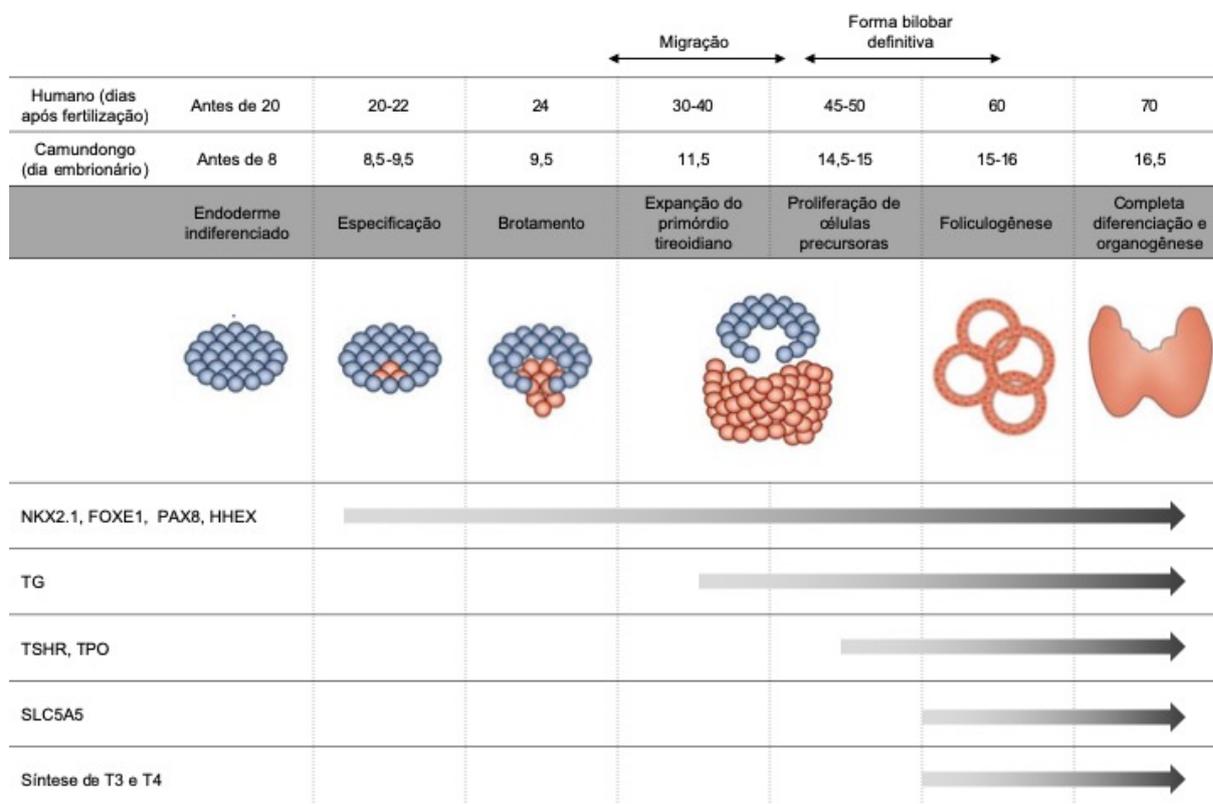


Figura 1: Fases do desenvolvimento embrionário da tireoide humana e de camundongo. A ilustração mostra a expressão dos fatores de transcrição na tireoide e os estágios do desenvolvimento tireoidiano. Adaptado de Fernandez et.al, 2015 (7).

Em modelo de rato foi observado que a embriogênese normal da tireoide é controlada pela expressão hierárquica de fatores de transcrição (FTs), que formam uma rede de regulação integrada (8, 9). As células precursoras do divertículo tireoidiano, caracterizam-se pela expressão dos FTs *Nkx2-1*, *Pax8*, *FoxE1* e *Hhex* (4). Foi proposto que a sobrevivência do primórdio tireoidiano é dependente da expressão de *FoxE1* e *Pax8* (8) e que sua migração depende principalmente da expressão de *FoxE1* (10). Já foi mostrado que a expressão ectópica de *Pax8* e *Nkx2-1* promovem a diferenciação de células tronco de rato em células foliculares, capazes de se organizar em folículos (11). Por outro lado, nem o *Tsh* nem seu receptor (*Tshr*) seriam necessários para o desenvolvimento e diferenciação da glândula (5).

Experimentos em modelo de zebra-fish mostraram que, durante o desenvolvimento inicial da tireoide, a mesoderme em volta do primórdio tireoidiano

seria uma fonte de indutores, entre eles Hand2 e Tbx1 que regulam a via de sinalização do Fgf necessária para o desenvolvimento da tireoide (12, 13).

1.2 Hipotireoidismo congênito

O hipotireoidismo congênito (HC) é um distúrbio metabólico sistêmico, onde a produção dos HT no período neonatal é insuficiente. Ocorre em cerca de 1:3.000 a 1:4.000 recém-nascidos, sendo uma causa de retardo mental que pode ser prevenida quando diagnosticada e tratada precocemente. O HC deve ser detectado na triagem neonatal, através do “teste do pezinho”, e o diagnóstico etiológico é determinado aos três anos de idade. O início do tratamento de HC deve ser iniciado o quanto antes para evitar prejuízos no desenvolvimento do sistema neuropsicomotor (14, 15).

O HC primário, acometendo diretamente a glândula tireoide, é caracterizado por níveis elevados de TSH em resposta a níveis reduzidos de HT, que é essencial para o desenvolvimento do cérebro durante a fase fetal e início de vida pós-natal, sendo assim, o hipotireoidismo não tratado vai causar déficits cognitivos e motores. A gravidade desses problemas no neurodesenvolvimento dependerá da gravidade e da duração do estado de hipotireoidismo (16).

A incidência do HC pode variar dependendo da etnia, consanguinidade e área geográfica. A maior prevalência ainda é em hispânicos e caucasianos, e mais frequente no sexo feminino. Com o passar dos anos, a incidência do HC vem aumentando devido à diminuição dos valores de corte do TSH neonatal, além disso, ainda existe uma grande migração de populações mais afetadas para áreas de baixa incidência, contribuindo para este aumento (17-19). Em 2013, a incidência de HC no Brasil variava de 1:2.595 a 1:4.795 nascidos vivos (20). Outros estudos brasileiros mostram uma incidência de HC de 1:2.876 em Santa Catarina (21), 1:4.632 em Tocantins (22) e 1:4.166 em Sergipe (23). Os valores mundiais diferem devido aos diferentes valores de corte utilizados por cada país, bem como diferenças étnicas, socioculturais e territoriais.

Os pacientes com HC primário e permanente são divididos em dois grupos: aqueles que apresentam erros inatos da biossíntese dos HT (disormonogênese) e respondem por cerca de 10-15% dos casos, e aqueles que apresentam defeitos no

desenvolvimento da tireoide (disgenesia), esse grupo responde por 80-85% dos casos (24).

A disormonogênese tireoidiana é uma doença causada por erros inatos em um dos mecanismos necessários para a síntese dos HTs. A herança é autossômica recessiva e muito frequentemente leva ao aumento da glândula (bócio), como consequência dos elevados níveis de TSH (25). O diagnóstico etiológico da disormonogênese é realizado através de testes clínico-laboratoriais (tireoglobulina sérica, teste de descarga de perclorato) e moleculares (26). A disormonogênese pode ocorrer devido a um defeito molecular em qualquer uma das etapas da síntese ou da secreção dos HTs (27). Já foram descritas mutações nos genes *SLC5A5* (28), *SLC26A4* (29), *DUOX2* e *DUOXA2* (30), *TPO* (31), *TG* (32), e *DEHAL-1* (33).

Os defeitos no desenvolvimento da tireoide (que serão aprofundados na próxima sessão) são, frequentemente, casos esporádicos, sendo mais comuns em meninas, hispânicos, portadores de doenças cardíacas congênitas e luxação congênita do quadril (34). Os fenótipos da disgenesia variam entre hipoplasia (cerca de 5% dos casos), ectopia (48-61% dos casos) e agenesia (15-33% dos casos), os aspectos morfológicos, porém, são oriundos de eventos moleculares distintos. Os FTs *NKX2.1*, *FOXE1*, *PAX8* e *HHEX*, presentes desde os primeiros estágios da embriogênese tireoidiana, o *TSHR* e os fatores de maturação *HOXA3* e *EYA1*, expressos durante os estágios mais avançados da organogênese, podem estar envolvidos (2, 35-37).

1.2.1 Disgenesia tireoidiana

A disgenesia tireoidiana (DT) constitui um grupo heterogêneo de defeitos de desenvolvimento da tireoide com importantes variações fenotípicas (38). A ectopia ocorre na maioria dos casos de DT (65%) por falhas na migração da glândula durante seu desenvolvimento, que podem deixar parte ou todo o tecido tireoidiano na base da língua ou ao longo do trajeto do canal tireoglossal ou, mais raramente, na orofaringe, região infrahiodea, laringe, traqueia, mediastino, esôfago, coração ou em linfonodos cervicais (39). A segunda forma mais frequente das DTs é a atireose (30%) e pode ser considerada a forma mais grave (40). Ainda não está estabelecido se as células foliculares após sua diferenciação inicial desaparecem por apoptose ou se não há

diferenciação inicial (2). A forma menos frequente de DT é a hipoplasia (5%), com a glândula *in situ* de tamanho reduzido que, quando é muito pequena, pode ser confundida com atireose. Os defeitos de desenvolvimento tireoidiano também podem provocar hemiagenesias, com o desenvolvimento de um único lóbulo, com ou sem o istmo (40).

Apesar de a maioria dos casos de DT ser esporádica, cerca de 2% apresentam padrão de herança familiar, com preponderância do comprometimento do sexo feminino sobre o masculino (38, 41). Em estudo com aproximadamente 2.500 pacientes franceses com HC e DT observou-se maior número de casos entre os familiares de 1º grau, o que é o esperado para uma doença esporádica. Por outro lado, não observou-se concordância no fenótipo de irmãos afetados (ex. um com ectopia e outro com atireose), sugerindo um possível evento genético comum responsável por dois fenótipos diferentes (42). Os resultados obtidos dos casos familiares sugerem que o padrão de herança poderia ser autossômico recessivo ou dominante com penetrância incompleta, ou que poderia haver heterogeneidade genética, ou até mesmo o envolvimento de novos genes. No entanto, estudos de gêmeos observaram elevada discordância para HC por DT tanto em gêmeos monozigóticos como em dizigóticos (41-44).

Até o momento, a patogenia da maioria dos casos das DTs não foi determinada, somente em aproximadamente 5% dos pacientes estudados foram identificadas mutações em genes envolvidos na organogênese da tireoide (45, 46). Os defeitos anormais do desenvolvimento tireoidiano já foram relacionados a variantes nos genes *HOXA3*, *NKX2-1*, *PAX8*, *FOXE1*, *HHEX*, *HES1*, *EYA1*, *TSHR*, *TUBB1*, *GLIS3*, *JAG1*, *NKX2-5*, *NTN1* e *CDCA8* (47).

Na procura pelas causas genéticas das DTs foram também investigados genes que não se expressam exclusivamente na tireoide e que poderiam afetar o desenvolvimento glandular. Mutações no gene *TBX1*, que se expressa no mesoderma do embrião, foram associadas à hemiagenesia e hipoplasia, porém eram casos de Síndrome de Di George (48). Mutações germinativas no gene *SHH*, que participa da formação do palato e poderia estar envolvido na regulação da lobulação da tireoide, e no gene *ISL1*, fator de transcrição que se expressa na tireoide e no coração durante a morfogênese, não foram localizadas em pacientes com DT (49, 50).

Os estudos mais recentes também mostram casos de DT associados às alterações em genes que participam da hormonogênese tireoidiana. O gene *SLC26A4*, relacionado à síndrome de Pendred, estava mutado numa família consanguínea com DT (51) e em outro trabalho o gene *TPO* estava mutado em casos de hipoplasia tireoidiana (52). Um outro grupo brasileiro realizou o sequenciamento completo do exoma de pacientes bem caracterizados com DT, identificaram diversas variantes no gene *DUOX2*, além de caracterizar variantes funcionais neste gene relacionadas à ectopias tireoidianas (53).

Thorwarth, *et al.* (54) realizaram o mapeamento de variações de número de cópias (CNVs) de casos com defeito do desenvolvimento e função tireoidiana. Os resultados mostraram que somente 8,75% dos pacientes com hipoplasia ou atireose apresentavam alterações cromossômicas estruturais (<1 Mb). Estes resultados sugerem que grandes aberrações cromossômicas não seriam a causa principal das DTs. No entanto, foi descrito um caso de um paciente com HC, hemiagenesia e algumas outras malformações portador de microduplicação do cromossomo 22q11.2 (55).

Diante da escassez de defeitos genéticos já relatados envolvendo o desenvolvimento da tireoide, outros mecanismos, como mutações somáticas precoces ou modificações epigenéticas são propostos para explicar a DT. Esses fatores poderiam gerar consequências para o crescimento, desenvolvimento e, posteriormente, para saúde de indivíduos afetados (56). Alguns estudos analisaram a relação de fatores ambientais com a malformação tireoidiana e em todos eles os autores não encontraram dados conclusivos, suportando apenas uma hipótese (57-59).

Assim, observamos que: mutações em diferentes genes podem estar associadas a um mesmo fenótipo de DT (ectopia, atireose ou hipoplasia); mutações em um mesmo gene podem estar associadas aos diferentes fenótipos; múltiplos defeitos podem interferir na morfogênese da tireoide e, conseqüentemente, genes ainda não descritos, que se expressam em tireoide ou em outros tecidos, podem estar associados às DTs. Portanto, é possível afirmar que as causas das disgenesias tireoidianas "ainda" são um enigma a ser resolvido.

Considerando que a maioria dos casos de HC deve-se aos erros no desenvolvimento da glândula tireoide durante a embriogênese, os estudos mais

antigos focavam principalmente na pesquisa de mutações por sequenciamento direto de genes sabidamente associados ao desenvolvimento da tireoide que se expressam na própria glândula ou em tecidos extratireoidianos, ou em análises de ligação com marcadores para esses mesmos genes, ou na detecção de CNVs. Essas pesquisas permitiram explicar somente 5% dos casos das DTs e não conseguiram definir um padrão de herança, a presença de penetrância e ou expressividade, ou se as diferentes formas de apresentação clínica das DTs têm uma origem genética comum ou é multigenica. Por isso, é necessário continuar investigando as bases moleculares desta doença.

1.3 Sequenciamento de nova geração

O sequenciamento de DNA desenvolvido por Sanger et al. (60) é uma abordagem muito utilizada até hoje, apesar da técnica apresentar limitações de tempo e custos relativamente elevados, principalmente quando pretende-se estudar grandes sequências de DNA ou até mesmo um genoma completo. Diante destes obstáculos, novas tecnologias de sequenciamento denominadas de “sequenciamento de nova geração” (*Next Generation Sequencing* - NGS) foram desenvolvidas. Estas plataformas são capazes de gerar informações sobre milhões de pares de bases em uma única corrida (61, 62).

Uma das metodologias é o sequenciamento global do exoma ou *whole-exome sequencing* (WES) (63), em que o DNA genômico é reduzido até obter amostras ricas em regiões codificantes de proteínas, os exons ou exoma, e em seguida realizar um sequenciamento de alto rendimento (64). A técnica é uma ferramenta molecular poderosa pois permite não somente identificar mutações (*missense*, *nonsense*, sítios de *splicing*, pequenas deleções/inserções) como genes de fusão e variação do número de cópias de segmentos genômicos (65). Com o WES é possível sequenciar aproximadamente 200.000 exons, 23.500 genes ou 30 milhões de nucleotídeos de uma dúzia de indivíduos em poucos dias (37Mb) (66). Comparado ao sequenciamento global do genoma, o WES tem uma relação custo-benefício superior. Com um mesmo custo, pode-se alcançar uma cobertura média de 200X, os resultados são gerados 5 vezes mais rápido e resulta em um menor volume de dados para análise (64, 67) .

Dentro das pesquisas sobre as causas genéticas que envolvem todo o complexo desenvolvimento da tireoide, pesquisadores puderam identificar novos genes e mecanismos associados analisando dados de exoma de pacientes com DT. Já foram descritas mutações no gene *CDC8 (BOREALIN)* em uma família consanguínea, afetando a adesão e migração das células em desenvolvimento (68); no gene *DUOX2* em pacientes com ectopia, alterando a interação da proteína com o seu fator de maturação *DUOXA2* (53); no gene *TRPC4AP*, surgindo como um novo gene candidato para CH primário (69); no gene *GLI3* (em heterozigose compostas) em irmão com hemiagenesia tireoidiana (70).

2 OBJETIVO

O objetivo deste projeto foi ampliar o estudo das bases moleculares do hipotireoidismo congênito causado por disgenesia tireoidiana, através da pesquisa de alterações gênicas com relevância clínica no exoma completo e em sequências regulatórias, utilizando sequenciamento de nova geração e análises *in silico*, em uma coorte de pacientes com disgenesia tireoidiana cuidadosamente caracterizada em termos bioquímicos e morfológicos.

2.1 Objetivos específicos

1. Sequenciar o exoma completo e sequências regulatórias de pacientes com HC por DT que apresentam ectopia, atireose e hipoplasia;
2. Identificar variantes gênicas (SNPs, pequenas deleções e inserções) e CNVs nestas amostras sequenciadas;
3. Selecionar alterações gênicas de interesse potencialmente associadas ao desenvolvimento tireoidiano;
4. Validar as alterações selecionadas nas amostras submetidas ao sequenciamento do exoma expandido, bem como em amostras de outros pacientes com HC por DT e de indivíduos controles normais (com função tireoidiana normal);
5. Realizar análise de predição do efeito de alterações não sinônimas nas proteínas;
6. Avaliar as correlações já descritas entre as variantes gênicas e fenótipos, visando verificar seu efeito;
7. Categorizar funcionalmente os genes afetados quanto às vias de sinalização e classes funcionais envolvidas.

3 MÉTODOS

3.1 Pacientes

A casuística deste estudo foi composta por 71 pacientes brasileiros de diferentes localidades. Deste total, 69 amostras de sangue de crianças diagnosticadas na triagem neonatal, que tiveram o diagnóstico confirmado após 3 anos de idade. De Salvador (BA), 45 casos em colaboração com o Dr. Helton Ramos (Universidade Federal da Bahia) e a Dr^a. Tatiana Amorim (APAE Salvador). De Botucatu (SP), 10 casos em colaboração com a Dr^a. Célia Nogueira (Universidade Estadual Paulista). De Marília (SP) e 14 casos em colaboração com o Dr José Sgarbi (Faculdade de Medicina de Marília). Outras duas amostras de tecidos ectópicos diagnosticados em idade adulta e coletados após cirurgia foram inclusas, em colaboração com o Dr Geraldo Medeiros Neto (Faculdade de Medicina da Universidade de São Paulo) e o Dr Celso Friguglietti (Hospital Santa Catarina) de São Paulo (SP).

Do total de pacientes, 17 com o diagnóstico etiológico de agenesia (23,9%), 1 caso de hemiagenesia (1,4%), 30 com hipoplasia (42,3%), 17 com ectopia (23,9%) e 6 sem diagnóstico definido (8,5%). Dentre eles, 38 do gênero feminino e 33 do gênero masculino. Contamos ainda com um total de 128 indivíduos controle sem doença tireoidiana, sendo 83 amostras de São Paulo (SP), 29 amostras de Salvador (BA) e 16 amostras de pais de pacientes de Marília (SP). Os pacientes e/ou seus responsáveis, assim como seus familiares foram esclarecidos sobre os exames realizados e a importância do estudo e assinaram o termo de consentimento livre e esclarecido. Esse projeto possui aprovação no comitê de ética (CAAE: 59240816.2.0000.5505) (anexo 1).

3.2 Sequenciamento

O DNA dos pacientes e controles foi extraído de sangue periférico seguindo protocolo padrão (71). O DNA de tecido dos dois pacientes adultos que tiveram as tireoides removidas foi extraído macerando o tecido congelado com nitrogênio líquido, acrescentado cerca de 1mL de tampão (100mM Tris HCL pH 8.5, 5M EDTA, 10%

SDS, 5M NaCl) e Proteinase K (0,2mg), seguida de incubação 37°C por 16 horas, as etapas seguintes foram feitas seguindo protocolo padrão.

3.2.1 Sequenciamento do exoma

Selecionamos 33 amostras provenientes da APAE Salvador e as 2 amostras de tecido para realizar o sequenciamento global do exoma e regiões regulatórias (WES), em parceria com a Dr^a. Anamaria Camargo do Centro de Oncologia Molecular do Instituto Sírio Libanês de Ensino e Pesquisa. Os dados clínicos e laboratoriais das 35 amostras estão no anexo 2.

A montagem das bibliotecas foi realizada com cerca de 200ng de DNA e utilizando o estojo comercial SureSelectXT Human All Exon V6+UTR Capture Library (Agilent Technologies, EUA), que enriquece as regiões codificadoras e UTR adjacentes. O sequenciamento foi realizado em plataforma Illumina NextSeq 500 (Illumina Inc., EUA) com o estojo comercial Illumina NextSeq 500 High Output V2 (300 cycles) (Illumina Inc., EUA) com cobertura média calculada e esperada para cada amostra de aproximadamente 140X.

O processamento dos dados de sequenciamento foi realizado em parceria com o Dr Pedro Galante, também do Centro de Oncologia Molecular do Instituto Sírio Libanês de Ensino e Pesquisa. Após o sequenciamento das bibliotecas, as amostras foram alinhadas com sequencia referência do genoma humano (GRCh37/hg19) utilizando a ferramenta Burrows-Wheeler Alignment (BWA-MEM) (72). Com os pacotes SAMTools e PicardTools foram aplicados filtro “q30” de qualidade para selecionar os melhores alinhamentos e métricas de avaliação do arquivo BAM, respectivamente. Os alinhamentos foram submetidos às ferramentas Genome Analysis Toolkit (GATK) (73), onde foi realizada a recalibração de bases, a chamada de variantes de nucleotídeos únicos e pequenas inserções e deleções e marcação de variantes com “hardfilter”. Por fim, as variantes identificadas foram anotadas de acordo com sua posição no genoma, utilizando a ferramenta ANNOVAR (74) e gerado o arquivo VCF. Foram mantidas as variantes com no mínimo de 2 leituras alteradas e marcadas com “pass” para SNVs e no mínimo 3 leituras alteradas para indels.

Ao fim do processamento obtivemos os dados completos do sequenciamento dos pacientes, contendo variantes de troca única (SNV – single nucleic variant) e

pequenas deleções e inserções (indel), identificando cada variante de acordo com sua posição cromossômica e ainda trazendo informações de diversos bancos de dados incluídos na anotação do ANNOVAR.

3.2.2 Sequenciamento Sanger

Para a confirmação de variantes pontuais nos pacientes e controles, utilizamos a técnica de sequenciamento pelo método de Sanger.

Cerca de 100ng de DNA foi amplificado com o estojo comercial GoTaq DNA Polymerase (Promega Corporation, Madison, Wisconsin, EUA) e primers específicos (anexo 3) desenhados com a ferramenta PrimerQuest (Integrated DNA Technologies, Inc., Coralville, Iowa, EUA). O sequenciamento foi realizado em equipamento ABI 377 com o estojo comercial BigDye Terminator v3.1 (Life Technologies, Carlsbad, California, EUA). As sequências foram analisadas com o programa Sequencher (Gene Codes Corporation, Ann Arbor, Michigan, EUA) utilizando a mesma sequência do genoma humano (GRCh37/hg19) como referência.

3.3 Análise dos dados de sequenciamento do exoma

Com os dados de sequenciamento do exoma de cada paciente, prosseguimos para a seleção e análise de alterações.

3.3.1 Análise de variantes

A seleção de variantes foi realizada a partir de uma sequência de filtragens específicas para cada análise de nosso interesse, que foram:

- a. Genes alterados em comum:** verificamos a presença de genes alterados em comum entre todos os pacientes com disgenesia e entre os grupos de fenótipo.
- b. Variantes em comum:** buscamos por variantes em comum entre todos os pacientes com disgenesia e entre os grupos de fenótipo.
- c. Variantes em genes candidatos:** buscamos por variantes presentes em 147 genes candidatos (anexo 5). Dentro desta análise, ainda avaliamos:

- **Genes regulados por fatores de transcrição:** buscamos por variantes presentes nos genes que são regulados pelos FTs PAX8 (anexo 6) e FOXE1 (anexo 7).
- **Análise do gene *FOXE1*:** realizamos uma análise detalhada das variantes presentes em toda a região promotora e codificadora do gene *FOXE1*.
- **Análise dos genes da família DUOX:** realizamos uma análise detalhada das variantes presentes nos genes *DUOX1*, *DUOXA1*, *DUOX2* e *DUOXA2*.

d. Regiões em homozigose: Investigamos grandes regiões em homozigose.

Para as análises a, b e c, selecionamos as variantes de nosso interesse seguindo os seguintes critérios: (a) variantes com a relação entre as frequências de leitura da base alterada e base referência maior que 20, removendo assim erros de leitura; (b) variantes nas regiões codificadora, promotora (5'UTR) e de *splicing* e (c) variantes não sinônimas.

As análises de frequência foram realizadas sempre selecionando variantes presentes em menos que 1% das populações (variantes raras). O critério para seleção de variantes deletérias foi a classificação como causadora de doenças em, pelo menos, dois programas de predição.

3.3.2 Análise de CNVs

Com o desenvolvimento de novas tecnologias é possível identificar CNVs em dados de sequenciamento. Utilizamos o programa XHMM (*eXome-Hidden Markov Model*) (28), considerando o Q-Exact (coeficiente de qualidade de CNV exata) >80%. Optou-se por um valor de corte bem elevado para evitar identificar falsas CNVs, pois o programa recomenda utilizar dados com elevado número de leituras.

3.4 Ferramentas de bioanálise

Para auxiliar na análise dos dados e trazer mais informações sobre os achados genéticos dos nossos pacientes, utilizamos diversos recursos online de bancos de dados e programas de bioinformática.

3.4.1 Bancos de dados

Utilizamos informações contidas em bancos de dados disponíveis na internet para coletar mais informações sobre a frequência populacional das variantes e fenótipos associados, ajudando a selecionar dados específicos de acordo com a análise realizada. São eles:

- *The 1000 Genomes Project* (1000G). Catálogo público de variantes humanas e informações de genótipos (75).
- Arquivo Brasileiro Online de Mutações (ABraOM). Repositório de variantes genômicas provenientes do sequenciamento do exoma de indivíduos brasileiros (76).
- ClinVar. Arquivo público de acesso livre de relatórios sobre as relações entre variações e fenótipos humanos, com evidências de apoio (77).
- *Database of Single Nucleotide Polymorphisms* (dbSNP). Arquivo público para variações de nucleotídeo único humano, microssatélites e inserções e exclusões em pequena escala (78).
- *Exome Aggregation Consortium* (ExAc). Coalizão de investigadores que buscam agregar e harmonizar dados de sequenciamento de exoma de uma variedade de projetos de sequenciamento em larga escala (79).
- *Database of Genomic Variants* (DGV) Catálogo de variantes genômicas estruturais humanas (80).

3.4.2 Programas de bioinformática

Utilizamos programas que fazem predições sobre o efeito estrutural e funcional de variantes em proteínas ou sítios de ligação de FTs, análise de enriquecimento de vias e interação de proteínas para a seleção e análise de variantes. São eles:

- *Enrichr*. ferramenta que avalia vias enriquecidas, mecanismos e outras informações sobre genes (82).

- *FannsDB-Condol: Functional Annotations for Non-Synonymous SNVs (Consensus Deleteriousness)*. Método que combina várias ferramentas para avaliar o efeito de SNVs não sinônimos (83).
- *FATHMM-MKL: Functional Analysis through Hidden Markov Models*. Prevê as consequências funcionais de SNVs não sinônimos e variantes não codificantes (84).
- *HOPE*: ferramenta que analisa o efeito estrutural de mutações na sequência proteica (85).
- *MutationAssessor: Functional Impact of Protein Mutations*. Ferramenta que prevê o impacto funcional da substituição de aminoácidos na proteína (86).
- *MutationTaster*. Ferramenta que prevê a consequência de variantes genômicas (87).
- *PolyPhen2: Polimorphism Phenotyping v2*. Ferramenta que prevê o possível impacto de SNVs na estrutura e função de proteínas (88).
- *String*: avalia enriquecimento de vias e interação de proteínas (89).
- *TFBind*: Software para pesquisa de sítios de ligação de FTs (90).

3.4.3 Análises estatísticas

Para a análise estatística dos dados utilizamos o software *GraphPad Prism* (GraphPad Prism version 6.0.0 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com). A associação de variantes e o desenvolvimento de DT foi analisada usando o teste *Pearson's Chi-square*. A análise de regressão logística foi utilizada para determinar a razão de possibilidade (*odds ratio* – OR) e intervalo de confiança de 95% (95%CI).

4 RESULTADOS

Os dados gerais do *WES*, bem como a qualidade da leitura das bases, o número total de bases sequenciadas, bases sem cobertura e a cobertura do sequenciamento estão apresentados no Anexo 4. Obtivemos um sequenciamento uniforme, de qualidade e semelhante para as 35 amostras. Uma média de 85.600,00 bases foram sequenciadas, sendo que apenas entre 500 e 900 bases não foram lidas para cada pacientes. Todas as amostras apresentaram mais de 63% das bases com cobertura maior ou igual à 30X.

4.1 Causas genéticas em comum

A primeira análise realizada nos dados do exoma dos pacientes foi verificar a ocorrência de causas genéticas comuns entre os pacientes. Avaliamos genes alterados em comum e variantes em comum.

4.1.1 Genes alterados em comum

Primeiro, verificamos quais genes estavam alterados em comum entre os pacientes, analisando tanto individualmente cada grupo de fenótipo (agenesias, hipoplasias e ectopias) quanto todo o conjunto de pacientes com disgenesia. Os genes foram selecionados independentemente do tipo e do número de variantes por gene.

Verificamos a presença de 4.565 genes alterados em comum, dentre eles: 3.622 genes alterados nas agenesias, 3.387 nas hipoplasias e 3.344 nas ectopias. Alguns destes genes se repetiam entre cada dois grupos de fenótipos e até mesmo dentro de todo o grupo das disgenesia. Os dados sobre a distribuição destes genes estão compilados na Figura 2. O diagrama de Venn mostra que do total de 4.565 genes, 2.393 (52,4%) estavam alterados em comum entre todo o grupo das disgenesia. Para ter uma ideia geral sobre a função sua função, submetemos estes 2.393 genes alterados em todas as disgenesias na ferramenta de análise *Enrichr* para verificar a presença de vias enriquecidas ou mecanismo associados aos mesmos. A análise gerada pelo programa *Enrichr* resultou em 285 termos associados aos genes

avaliados e selecionamos exclusivamente aqueles relacionados ao desenvolvimento. Esta seleção resultou em 38 termos associados aos eventos de diferenciação, proliferação, migração, adesão e morfogênese celular (Tabela 1).

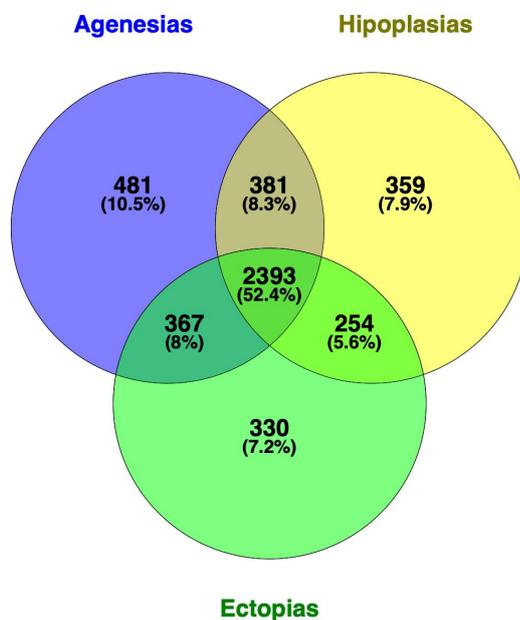


Figura 2: Diagrama de Venn mostrando o número de genes alterados no exoma das 35 amostras de DT sequenciadas. Genes alterados nas agenesias: 3.622, hipoplasias: 3.387, ectopias: 3.344. Genes alterados exclusivamente nas agenesias: 481, hipoplasias: 359, ectopias: 330. Genes alterados entre agenesias e hipoplasias: 381, entre hipoplasias e ectopias: 254, entre ectopias e agenesias: 367. Em comum entre todas as disgenesias: 2.393. Total de genes alterados: 4.565.

Tabela 1: Termos selecionados resultantes da pesquisa feita dentro da ferramenta *Enrichr* dos 2.393 genes alterados em comum no exoma das 35 amostras de DT sequenciadas (termos com $p < 0,05$) (continua)

Fonte	Termo	ID
KEGG	<i>Focal adhesion</i>	map04510
GO_BP	<i>Cell morphogenesis involved in differentiation</i>	GO:0000904
GO_BP	<i>Epithelial cell migration</i>	GO:0010631
GO_BP	<i>ERBB signaling pathway</i>	GO:0038127
GO_BP	<i>ERBB2 signaling pathway</i>	GO:0038128
GO_BP	<i>Negative regulation of cell migration</i>	GO:0030336
GO_BP	<i>Negative regulation of cell motility</i>	GO:2000146
GO_BP	<i>Negative regulation of dendrite development</i>	GO:2000171
GO_BP	<i>Negative regulation of nervous system development</i>	GO:0051961
GO_BP	<i>Negative regulation of Notch signaling pathway</i>	GO:0045746
GO_BP	<i>Photoreceptor cell development</i>	GO:0042461

Tabela 1: Termos selecionados resultantes da pesquisa feita dentro da ferramenta *Enrichr* dos 2.393 genes alterados em comum no exoma das 35 amostras de DT sequenciadas (termos com $p < 0,05$) (conclusão)

Fonte	Termo	ID
GO_BP	<i>Positive regulation of cell migration</i>	GO:0030335
GO_BP	<i>Positive regulation of cell migration by vascular endothelial growth factor signaling pathway</i>	GO:0038089
GO_BP	<i>Positive regulation of cell morphogenesis involved in differentiation</i>	GO:0010770
GO_BP	<i>Positive regulation of cell motility</i>	GO:2000147
GO_BP	<i>Positive regulation of cell-matrix adhesion</i>	GO:0001954
GO_BP	<i>Positive regulation of dendrite development</i>	GO:1900006
GO_BP	<i>Positive regulation of dendritic spine morphogenesis</i>	GO:0061003
GO_BP	<i>Positive regulation of endothelial cell migration</i>	GO:0010595
GO_BP	<i>Positive regulation of epithelial cell migration</i>	GO:0010634
GO_BP	<i>Positive regulation of ERBB signaling pathway</i>	GO:1901186
GO_BP	<i>Positive regulation of focal adhesion assembly</i>	GO:0051894
GO_BP	<i>Positive regulation of substrate adhesion-dependent cell spreading</i>	GO:1900026
GO_BP	<i>Regulation of cell migration</i>	GO:0030334
GO_BP	<i>Regulation of cell morphogenesis</i>	GO:0022604
GO_BP	<i>Regulation of chemokine-mediated signaling pathway</i>	GO:0070099
GO_BP	<i>Regulation of dendrite development</i>	GO:0050773
GO_BP	<i>Regulation of dendrite morphogenesis</i>	GO:0048814
GO_BP	<i>Regulation of dendritic spine morphogenesis</i>	GO:0061001
GO_BP	<i>Regulation of focal adhesion assembly</i>	GO:0051893
GO_BP	<i>Regulation of homotypic cell-cell adhesion</i>	GO:0034110
GO_BP	<i>Regulation of lamellipodium morphogenesis</i>	GO:2000392
GO_BP	<i>Regulation of macrophage derived foam cell differentiation</i>	GO:0010743
GO_BP	<i>Regulation of mammary gland epithelial cell proliferation</i>	GO:0033599
GO_BP	<i>Regulation of osteoclast development</i>	GO:2001204
GO_BP	<i>Regulation of substrate adhesion-dependent cell spreading</i>	GO:1900024
GO_BP	<i>Substrate adhesion-dependent cell spreading</i>	GO:0034446
GO_BP	<i>Substrate-dependent cell migration</i>	GO:0006929

KEGG: Kyoto Encyclopedia of Genes and Genomes, 2019; GO_BP: Gene Ontology, Biological Process, 2018.

Esta análise sugere o compartilhamento de genes com variantes associados a desenvolvimento, morfogênese, migração e adesão celular, não sendo possível definir algum padrão de genes alterados em comum entre todas as disgenesias.

4.1.2 Variantes em comum

Buscamos, em seguida, por variantes em comum entre os pacientes, analisando tanto cada grupo de fenótipo (agenesias, hipoplasias e ectopias) quanto todo o conjunto de pacientes com disgenesia. Esta análise mostrou a presença de 6.830 alterações em comum, dentre elas: 4.223 variantes nas agenesias, 1.602 nas hipoplasias e 1.699 nas ectopias. Algumas destas variantes se repetiam entre cada dois grupos de fenótipos e até mesmo dentro de todo o grupo das disgenesia. Os dados sobre a distribuição destas variantes estão compilados na Figura 3.

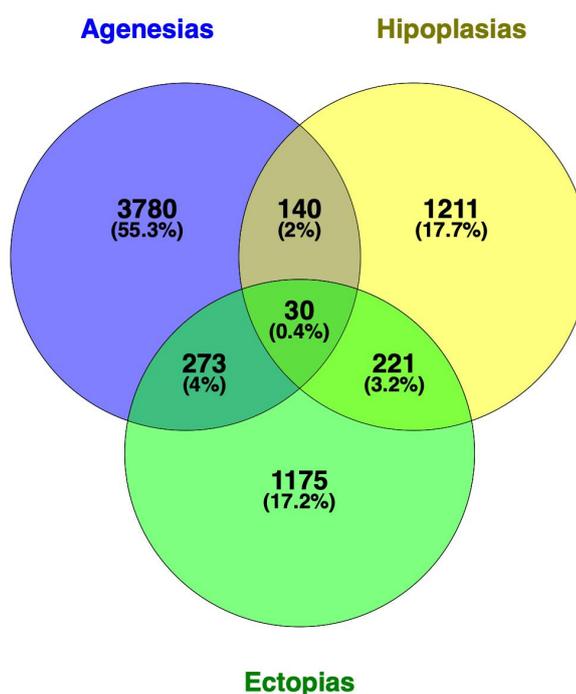


Figura 3: Diagrama de Venn mostrando o número de variantes em comum no exoma das 35 amostras de DT sequenciadas. Variantes nas agenesias: 4.223, hipoplasias: 1.602, ectopias: 1.699. Variantes exclusivamente nas agenesias: 3.780, hipoplasias: 1.211, ectopias: 1.175. Variantes em comum entre agenesias e hipoplasias: 140, entre hipoplasias e ectopias: 221, entre ectopias e agenesias: 273. Em comum entre todas disgenesias: 30. Total de variantes: 6.830.

Ao total de 6.830 variantes em comum foram aplicados filtros de frequência com dados populacionais dos bancos *ExAc*, *1000G* e *ABraOM*, selecionando as variantes raras (frequência menor que 1%). Também selecionamos as variantes deletérias com os dados anotados das predições geradas pelas ferramentas *FathMM-MKL*, *PolyPhen2*, *Condel* e *MutationTaster*. Esta análise resultou em 31 variantes em comum presentes em 7 genes (Figura 4 e Tabela 2).

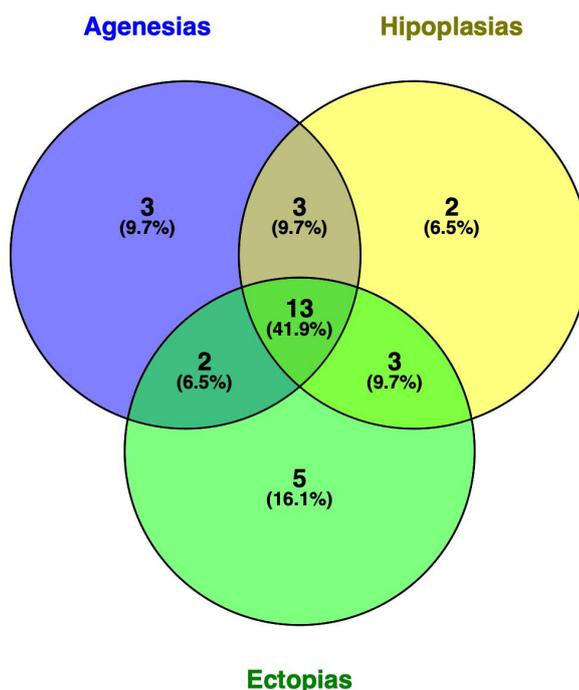


Figura 4: Diagrama de Venn mostrando as 31 variantes raras e deletórias em comum no exoma das 35 amostras de DT sequenciadas. Variantes nas agenesias: 21, hipoplasias: 21, ectopias: 23. Variantes exclusivamente nas agenesias: 3, hipoplasias: 2, ectopias: 5. Variantes em comum entre agenesias e hipoplasias: 3, entre hipoplasias e ectopias: 3, entre ectopias e agenesias: 2. Em comum entre todas as disgenesias: 13.

Tabela 2: Variantes raras e deletórias em comum identificadas no exoma das 35 amostras de DT sequenciadas (continua)

Fenótipo			Gene_variante	Função	dbSNP	Predição			
AG	HP	EC				FAMM	COND	PPH2	MT
	•		<i>CDC27_p.N575K</i>	NS	rs62075617	D	D	D	D
	•	•	<i>CDC27_p.K527Q</i>	NS	rs62075620	D	D	D	D
		•	<i>CDC27_p.S517G</i>	NS	rs62075623	D	D	D	D
	•	•	<i>CDC27_p.I306M</i>	NS	rs77739281	D	N	N	D
		•	<i>CDC27_p.L214V</i>	NS	rs62077263	D	N	D	D
	•	•	<i>CDC27_p.N199H</i>	NS	rs3208659	D	N	D	D
	•		<i>CDC27_p.P120S</i>	NS	rs149474782	D	N	N	D
	•	•	<i>CTDSP2_p.I106T</i>	NS	rs76940645	D	D	D	D
	•	•	<i>CTDSP2_p.V104M</i>	NS	rs111346934	D	N	D	D
	•	•	<i>CTDSP2_p.D98N</i>	NS	rs74343811	D	N	D	D
		•	<i>FRG2C_c.462_463T</i>	FS	rs144577984	-	-	-	D
	•	•	<i>HYDIN_p.E4160Q</i>	NS	rs1798314	D	N	D	N
	•	•	<i>HYDIN_c.11712_11713C</i>	FS	rs565285670	-	-	-	D
	•	•	<i>HYDIN_p.V3742I</i>	NS	rs1798413	D	N	D	N
	•	•	<i>HYDIN_p.G2558E</i>	NS	rs8044142	D	N	D	N

Tabela 2: Variantes raras e deletérias em comum identificadas no exoma das 35 amostras de DT sequenciadas (conclusão)

Fenótipo			Gene_variante	Função	dbSNP	Predição			
AG	HP	EC				FAMM	COND	PPH2	MT
•	•	•	<i>HYDIN</i> _c.7561_7576C	NFS	rs375727122	-	-	-	-
•	•	•	<i>HYDIN</i> _p.R2298G	NS	rs1774360	D	N	D	D
•	•	•	<i>HYDIN</i> _p.R2087C	NS	rs1774541	D	N	D	D
•	•	•	<i>HYDIN</i> _p.N751D	NS	rs3817211	D	N	D	D
•	•	•	<i>HYDIN</i> _p.T717A	NS	rs10744982	D	N	D	D
			<i>HYDIN</i> _p.G516D	NS	rs62040318	D	N	D	D
			<i>IGSF3</i> _p.I619T	NS	rs138851517	D	N	N	D
			<i>IGSF3</i> _p.W575X	SG	rs61730489	D	-	-	D
			<i>IGSF3</i> _p.R456C	NS	rs61786577	D	N	D	D
			<i>IGSF3</i> _p.D254N	NS	rs61786651	D	N	D	D
			<i>SETD8</i> _p.C302R	NS	rs61955126	D	N	N	D
•			<i>ZNF717</i> _p.H832Q	NS	rs79811623	N	D	D	N
•			<i>ZNF717</i> _p.L829P	NS	rs76707683	N	D	D	N
•			<i>ZNF717</i> _c.2010delinsAG	FS	rs139062880	-	-	-	D
•	•		<i>ZNF717</i> _p.C649X	SG	rs79635065	N	-	-	D
•			<i>ZNF717</i> _c.1127_1129C	FS	rs141065192	-	-	-	D

AG: agenesia; HP: hipoplasia; EC: ectopia; NS: troca de único nucleotídeo não sinônima; FS: indel causadora de frameshift; NFS: indel não causadora de frameshift; SG: stopgain; FAMM: *Fathmm-MKL*; COND: *Condel*; PPH2: *PolyPhen2*; MT: *MutationTaster*; D: deletério; N: Neutro.

Em seguida, buscamos por dados na literatura sobre cada uma dessas 31 variantes raras e deletérias selecionadas, com o objetivo de reunir informações sobre em qual quadro clínico já foram descritas, se seriam polimorfismos ou mutações, ou se estariam envolvidas com o desenvolvimento. Entretanto, nossas buscas mostraram que nenhuma delas é descrita na literatura junto à dados informativos sobre o quadro clínico dos pacientes em que foram identificadas. Submetemos a lista de 7 genes contendo as 30 variantes raras e deletérias na ferramenta de análise *Enrichr* para verificar a presença de vias enriquecidas ou mecanismo associados aos genes. A análise gerada pelo programa *Enrichr* resultou em 38 termos associados aos genes avaliados e selecionamos exclusivamente aqueles relacionados ao desenvolvimento. Esta seleção resultou em 3 termos (Tabela 3).

Tabela 3: Termos selecionados resultantes da pesquisa feita dentro da ferramenta *Enrichr* dos 7 genes com variantes raras e deletérias em comum identificadas no exoma das 35 amostras de DT sequenciadas (termos com $p < 0,05$)

Fonte	Termo	Genes
KEGG	Cell cycle	<i>CDC27</i>
GO_BP	exocrine system development (GO:0035272)	<i>IGSF3</i>
GO_BP	gland development (GO:0048732)	<i>IGSF3</i>

KEGG: Kyoto encyclopedia of genes and genomes 2019; GO_BP: Gene ontology biological process 2018.

Os termos selecionados na tabela 3 são relacionados aos genes *CDC27* e *IGSF3*. O *CDC27* é um componente do complexo promotor de anáfase altamente conservado em células eucariontes. A proteína interage com proteínas de checkpoint mitótico e pode estar envolvida no controle do tempo de mitose. O gene foi associado ao carcinoma de colo de útero e tularemia glandular. O gene apresenta múltiplos transcritos e possui pseudogenes já identificados nos cromossomos 2, 22 e Y (*RefSeq*, Mai 2014). Verificamos que *CDC27* está envolvido na via de ciclo celular e variantes neste gene foram identificadas em todas as amostras sequenciadas. O gene *IGSF3* codifica uma proteína de membrana do tipo imunoglobulina, mutações neste gene causam obstrução do ducto nasolacrimal bilateral (*RefSeq*, Jun 2016). O *IGSF3* também foi descrito na ativação de *USP28* que leva à apoptose em alguns tipos de câncer (91). O *IGSF3* aparece em anotações no *Gene Ontology* para o desenvolvimento do sistema exócrino e para o desenvolvimento de glândulas. Identificamos variantes neste gene em todos os casos de hipoplasia e ectopia. Também utilizamos a ferramenta *String* para verificarmos a interação dos 7 genes contendo variantes raras e deletérias em comum. No *String*, as interações entre proteínas são pontuadas e indicam confiança, baseado em previsão computacional, transferência de conhecimento entre organismos e interações agregadas de outros bancos de dados. Todas as pontuações variam de 0 a 1, sendo 1 a maior confiança possível. Pontuações de alta confiança ($>0,7$) mostram que as proteínas participariam de um mesmo processo biológico (89). O esquema de interação está representado na Figura 5, onde observamos duas interações de média confiança: entre os genes *CDC27* e *ZNF717* e entre os genes *ZNF717* e *FRG2C*. Os genes *SETD8*, *HYDIN*, *IGSF3* e *CTDSP2* não apresentaram interações.

Nesta etapa da análise identificamos variantes raras e deletérias em comum nos pacientes em genes que, inicialmente, não parecem ter relação com o

desenvolvimento e função tireoidiana. Uma análise detalhada futura, incluindo análises *in silico* e até ensaios funcionais, destas variantes poderá trazer mais informações sobre a real associação destes genes na causa da DT.

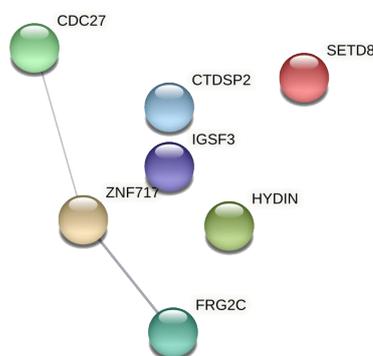


Figura 5: Interação dos 7 genes com variantes raras e deletérias em comum identificadas no exoma das 35 amostras de DT sequenciadas. Imagem gerada pela ferramenta *String*. Pontuação das interações: $ZNF717_FRG2C = 0,600$; $CDC27_ZNF717 = 0,403$. Pontuação das interações representam o grau de confiança: baixa (0,0 – 0,3), média (0,4 – 0,6) ou alta (0,7 – 1,0).

4.2 Variantes em genes candidatos

A seguir, estreitamos as análises dos dados de exoma focando agora em genes diretamente relacionados à tireoide em cada um dos pacientes. Para isso, compilamos uma lista contendo genes que conhecidamente participam da função e desenvolvimento tireoidiano, bem como genes diferentemente expressos no tecido. Chegamos em 147 genes, e chamamos estes de genes candidatos (Anexo 5).

Buscamos por alterações em toda a região codificadora e promotora destes genes, e chegamos ao total de 750 variantes em 115 genes. Aplicamos filtros de frequência com dados populacionais dos bancos *ExAc*, *1000G* e *ABraOM*, selecionando as variantes raras (frequência menor que 1%). Com as ferramentas *FathMM-MKL*, *MutationTaster*, *Condel* e *PolyPhen2* anotamos a predição das variantes e selecionamos aquelas com potencial deletério. Informações sobre as variantes contidas no banco de dados *ClinVar* também foram anotados. Esta análise resultou em 61 variantes raras e deletérias presentes em 42 genes (Tabela 4).

Tabela 4: Variantes raras e deletérias em genes candidatos identificadas no exoma das 35 amostras de DT sequenciadas (continua)

Pt	Gene variante	Função	dbSNP	ClinVar	Predição			
					FAMM	COND	PPH2	MT
	GLI2_p.S1555P	NS	rs144372453	B	D	N	D	D
1	<i>KDR_p.T771M</i>	NS	rs149745504	PB	D	D	D	D
	<i>TPO_p.P379R</i>	NS	rs575126367	-	N	N	D	D
2	<i>TG_p.S1506N</i>	NS	-	-	D	N	D	N
	<i>GLI2_p.Q1176H</i>	NS	rs139686081	B	D	N	D	D
3	<i>ISL1_p.N252S</i>	NS	rs121912286	-	D	N	N	D
	<i>PAX3_p.K183R</i>	NS	-	-	D	N	N	D
5	<i>CHGA_p.S98N</i>	NS	rs77938104	-	D	N	D	D
6	<i>DUOX2_p.T423I</i>	NS	rs201197899	-	D	D	N	D
7	<i>PRKCE_c.-24C>G</i>	RP	rs200186038	-	D	-	-	D
	<i>ABCA13_p.Q2357H</i>	NS	-	-	N	D	D	N
	<i>EGR1_p.P218L</i>	NS	-	-	D	N	D	D
9	<i>TCAP_p.S64L</i>	NS	rs45458802	B	D	N	D	D
	<i>TNXB_p.F1701L</i>	NS	-	I	N	D	D	N
10	BMP4_p.R287H	NS	rs121912768	B	D	D	N	D
12	DUOX2_p.F966SfsX29	FS	rs530719719	P	-	-	-	D
	<i>PITX2_c.-414C>G</i>	RP	-	-	D	-	-	D
	<i>BCL2L1_p.E139K</i>	NS	rs751979681	-	D	N	N	D
13	<i>TXNRD1_p.D82E</i>	NS	rs368438379	-	D	D	D	D
	GLI2_p.R720H	NS	rs149091975	I	D	N	D	D
14	KDR_p.G539R	NS	rs55716939	PB	D	N	N	D
	<i>TNXB_p.V906L^c</i>	NS	-	-	D	N	D	N
	<i>CREB1_p.T35A</i>	NS	rs114514847	-	D	D	N	D
15	<i>ISL1_p.P146T</i>	NS	rs200172777	-	D	D	D	D
	BGN_p.E104K	NS	rs150272427	PB	D	N	D	D
16	<i>DUOXA1_p.T168M^b</i>	NS	rs149960164	-	D	D	D	D
	<i>PBX4_p.G251S</i>	NS	rs200171479	-	D	N	D	D
19	<i>ARNT_p.D496N</i>	NS	rs1805133	B	D	N	N	D
	<i>FGF3_p.K101N*</i>	NS	rs61623544	B	D	N	D	D
20	<i>TSHR_p.T574S</i>	NS	rs61742289	PB	D	N	D	D
21	<i>ABCA13_p.S674F</i>	NS	rs189826199	-	D	D	D	D
22	<i>EYA1_c.-6750G>T</i>	RP	-	-	D	-	-	D
	<i>COL1A1_c.-57G>A</i>	RP	rs2734278	-	D	-	-	D
23	<i>ENO3_p.K5N</i>	NS	-	-	D	N	N	D
24	<i>MYBPC1_p.R535W</i>	NS	rs12146878	-	D	N	D	D

Tabela 4: Variantes raras e deletérias em genes candidatos identificadas no exoma das 35 amostras de DT sequenciadas (conclusão)

Pt	Gene_variante	Função	dbSNP	ClinVar	Predição			
					FAMM	COND	PPH2	MT
	<i>ABCA13</i> _p.T4713N	NS	rs193229571	-	D	N	D	D
25	<i>ACTA1</i> _c.-4C>T	RP	rs186826063	PB	D	-	-	D
	<i>COL1A1</i> _p.P823A ^a	NS	rs1800214	I	D	D	N	D
26	<i>HES1</i> _c.-213G>T	RP	-	-	D	-	-	D
27	<i>LEFTY1</i> _p.G312R	NS	rs143958321	-	D	N	D	D
	<i>CKM</i> _c.-3117G>C	RP	-	-	D	-	-	D
	<i>CKM</i> _c.-3118A>C	RP	-	-	D	-	-	D
28	<i>HADHA</i> _p.R208H	NS	-	-	D	D	N	D
	<i>ISL1</i> _c.-214C>A	RP	rs532316340	-	D	-	-	D
	<i>TNXB</i> _p.V906L ^c	NS	-	-	D	N	D	N
	<i>ABCA13</i> _p.S3715L	NS	rs202145037	PB	D	D	D	D
29	<i>COL1A1</i> _p.N1222S	NS	rs751784955	-	D	D	N	D
	<i>DUOXA1</i> _p.T168M ^b	NS	rs149960164	-	D	D	D	D
	<i>GPNMB</i> _p.N197K	NS	rs17147995	-	D	N	D	D
30	<i>TG</i> _p.G1195R	NS	rs138469414	I	D	D	D	D
	<i>COL1A1</i> _p.P823A ^a	NS	rs1800214	I	D	D	N	D
	<i>FGF3</i> _p.R171H	NS	-	-	D	D	D	D
31	<i>FOS</i> _p.P233A	NS	rs143816008	PB	D	N	N	D
	<i>PAX2</i> _c.-427C>T	RP	rs575017530	-	D	-	-	D
	<i>TCAP</i> _p.A118V	NS	rs143233087	I	D	N	D	D
	<i>TNXB</i> _p.V906L ^c	NS	-	-	D	N	D	N
	<i>DUOX2</i> _p.R516H	NS	rs143818717	-	D	D	D	D
	<i>FAU</i> _c.-332C>G	RP	rs147538389	-	D	-	-	D
32	<i>FOSB</i> _p.T63N	NS	-	-	D	D	D	D
	<i>HOXA2</i> _c.-55A>G	RP	-	-	D	-	-	D
	<i>HOXA3</i> _p.P135S	NS	rs148350002	-	D	N	-	D
33	<i>BMP4</i> _p.R226W	NS	rs140590144	I	D	N	D	D
	<i>JAG1</i> _c.-222A>C	RP	rs756190529	-	D	-	-	D
34	<i>DUOXA1</i> _p.T168M ^b	NS	rs149960164	-	D	D	D	D
	<i>TG</i>_p.Q870H	NS	rs2229843	-	D	D	D	N
35	<i>SFRP2</i> _p.S235X	SG	-	-	D	-	-	D

Pt: paciente; *: variante homocigota; a, b e c: variantes que se repetem; NS: troca de único nucleotídeo não sinônima; FS: indel causadora de frameshift; SG: stopgain; RP: região promotora; B: benigno; PB: provavelmente benigno; P: patogênico; I: indeterminado; NF: não fornecido; FAMM: *Fathmm-MKL*; COND: *Condel*; PPH2: *PolyPhen2*; MT: *MutationTaster*; D: deletério; N: Neutro; em negrito as variantes com fenótipo associado.

Das 61 variantes, apenas uma foi identificada em homozigose (*FGF3*_p.K101N) e 3 variantes aparecem em comum entre 2 ou 3 pacientes (*COL1A1*_p.P823A, *DUOXA1*_p.T168M e *TNXB*_p.V906L). Também observamos que 30 genes aparecem alterados apenas uma vez, 6 genes alterados duas vezes, 5 genes alterados 3 vezes e 1 gene alterado 4 vezes. Não houve alguma prevalência relevante de genes alterados e não observamos genes com mais de uma alteração no mesmo paciente.

Em seguida, verificamos dados na literatura sobre cada uma das 61 variantes selecionadas, buscando informações sobre em qual quadro clínico foram descritas, se seriam de fato polimorfismos ou mutações, ou se estariam envolvidas com o desenvolvimento. As variantes *GLI2*_p.S1555P e *BMP4*_p.R287H já foram descritas em pacientes com malformação crânio facial (92, 93) e a variante *GLI2*_p.R720H foi identificada em casos de hipopituitarismo (94). Já as variantes *DUOX2*_p.F966SfsX29 e *TG*_p.Q870H foram associadas à casos de defeito na organificação do iodeto (95-97) e bócio (98, 99), respectivamente. Outras duas variantes foram associadas ao câncer de tireoide: *KDR*_p.G539R (100) e *BGN*_p.E104K (identificada em amostras de câncer de tireoide na base de dados TCGA).

Utilizamos a ferramenta *String* e verificamos a interação dos genes candidatos alterados contendo variantes raras e deletérias. Esta análise foi feita avaliando a interação de genes alterados individualmente por paciente. Observamos interações nos genes candidatos alterados dos pacientes 3, 25, 31, 32 e 33. Todas as interações mostram confiança média, com pontuação entre 0,443 e 0,672 (Figura 6).

Analisando as variantes presentes nos genes candidatos observamos o mesmo padrão de aleatoriedade descrito na literatura, onde genes conhecidamente envolvidos na organogênese tireoidiana são pouco alterados (47).

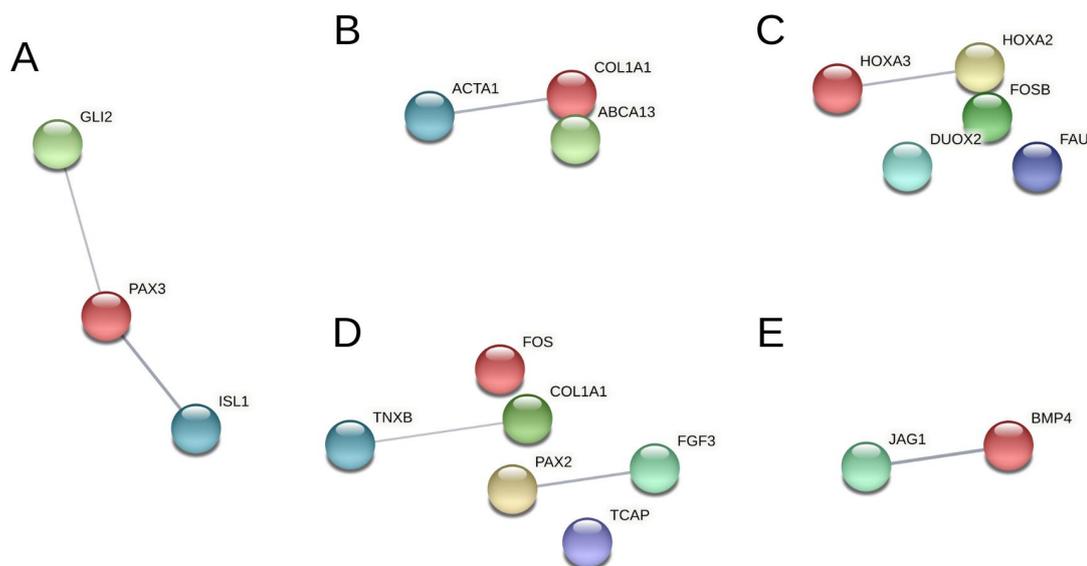


Figura 6: Interação dos genes candidatos alterados (análise por paciente) gerada pela ferramenta *String*. A. Paciente 3, pontuação das interações (PI): $PAX3_ISL1=0,639$ e $GLI2_PAX3=0,461$; B. Paciente 25, PI: $ACTA1_COL1A1=0,560$; C. Paciente 31, PI: $PAX2_FGF3=0,565$ e $TNXB_COL1A1=0,433$; D. Paciente 32, PI: $HOXA3_HOXA2=0,531$; E. Paciente 33, PI: $JAG1_BMP4=0,664$. Pontuação das interações representam o grau de confiança: baixa (0,0 – 0,3), média (0,4 – 0,6) ou alta (0,7 – 1,0).

4.2.1 Análises *in silico* de variantes em genes candidatos

Algumas variantes resultantes da análise anterior de genes candidatos se destacaram e decidimos explorar possíveis efeitos deletérios destas variantes através de análises *in silico*.

TSHR_p.T574S

Dentre os genes clássicos sabidamente envolvidos na DT, identificamos uma única alteração rara e deletéria no gene *TSHR*. A variante *TSHR_p.T574S* identificada em heterozigose no paciente 20 não é descrita na literatura associada a fenótipos normais ou alterados, apenas dados sobre frequência foram encontrados na base do *dbSNP*. Confirmamos por sequenciamento Sanger a presença da troca no paciente e não identificamos a alteração em 83 controles brasileiros normais.

A alteração se localiza na porção que compreende o domínio topológico extracelular entre os domínios transmembrana 4 e 5. A troca foi classificada com baixo impacto funcional na análise da ferramenta *MutationAssessor*. No entanto, a análise

gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e mostra sítios de *splicing* que podem ser perdidos ou ganhos, além de importantes domínios na proteína que podem ser perdidos devido a presença da variante. Observamos também que a região é bem conservada entre as espécies (Figura 7). A alteração se localiza na porção que compreende o domínio topológico extracelular entre os domínios transmembrana 4 e 5.

Gene TSHR humano NM_000369	Proteína cDNA gDNA	p.T574S c.1721C>G g.188791C>G	
Sítios de <i>splicing</i>			
Efeito	Posição	Pontuação	Borda exon-intron
Aceptor marginalmente aumentado	g.188787	wt: 0,5517 / mu: 0,6443	atgg ACAC
Aceptor marginalmente aumentado	g.188783	wt: 0,8943 / mu: 0,9306	gccc ATGG
Ganho de doador	g.188790	mu: 0,8800	GACA gcga
Características da proteína			
Início (aa)	Final (aa)	Característica	Efeito
561	580	Domínio topológico extracelular	Perda
581	602	Domínio transmembrana 5	Pode ser perdido
601	601	Conflito Y > H	Pode ser perdido
603	625	Domínio topológico citoplasmático	Pode ser perdido
626	649	Domínio transmembrana 6	Pode ser perdido
635	635	Conflito I > T	Pode ser perdido
645	645	Conflito L > V	Pode ser perdido
650	660	Domínio topológico extracelular	Pode ser perdido
661	682	Domínio transmembrana 7	Pode ser perdido
669	669	Conflito L > I	Pode ser perdido
683	764	Domínio topológico citoplasmático	Pode ser perdido
744	744	Conflito N > K	Pode ser perdido
762	764	MOTIF de ligação PDZ	Pode ser perdido
Conservação			
Espécies	Combinação	Alinhamento	
Humano	Referência	Y A K V S I C L P M D T E T P L A L A Y I V F	
Alterado	Conservada	Y A K V S I C L P M D S E T P L A L A Y I V F	
<i>P. troglodytes</i>	Indêntica	Y A K V S I C L P M D T E T P L A L A Y I V F	
<i>M. mulatta</i>	Indêntica	Y A K V S I C L P M D T E T P L A L A Y I V F	
<i>F. catus</i>	Indêntica	Y A K V S I C L P M D T E T P L A L A Y I I L	
<i>M. musculus</i>	Indêntica	Y A K V S I C L P M D T D T P L A L A Y I V L	
<i>G. gallus</i>	Indêntica	Y S K V S I C L P M D T E T P V A E A Y V V F	
<i>T. rubripes</i>	Indêntica	Y Q K V S I C L P M D T Q S I A A Q V Y I L S	
<i>D. rerio</i>	Indêntica	Y Q K V S I C L P M D T Q T L V D Q I Y I I C	
<i>D. melanogaster</i>	Não conservada	Y S S T S I C L P M E N R D V Y D T I Y L I A	
<i>C. elegans</i>	Não conservada	Y S E S S V C L P L R A A T I F D K S Y L I F	
<i>X. tropicalis</i>	Não homóloga		

Figura 7: Quadro informativo com dados gerados pela ferramenta *MutationTaster* sobre a variante p.T574S no gene *TSHR*. wt: *wild type* – selvagem; mu: mutado; Pontuação dos sítios de *splicing*: valores maiores que 0,3 representam modificação dos sítios.

Análise gerada pela ferramenta *HOPE* prediz diversas consequências que podem ser causadas pela troca:

- **Aminoácidos.** Os aminoácidos de tipo selvagem e mutantes diferem em tamanho. O resíduo mutante é menor, o que pode levar à perda de interações (Figura 8).
- **Conservação.** O resíduo de tipo selvagem não é conservado nesta posição. O resíduo mutante não estava entre os tipos de resíduo observados nesta posição em outras sequências homólogas, o que pode indicar que a mutação possivelmente está danificando a proteína. O resíduo mutante está localizado perto de uma posição altamente conservada.
- **Domínios.** O resíduo mutado está localizado em um domínio importante para a atividade principal da proteína. A mutação do resíduo pode perturbar essa função.

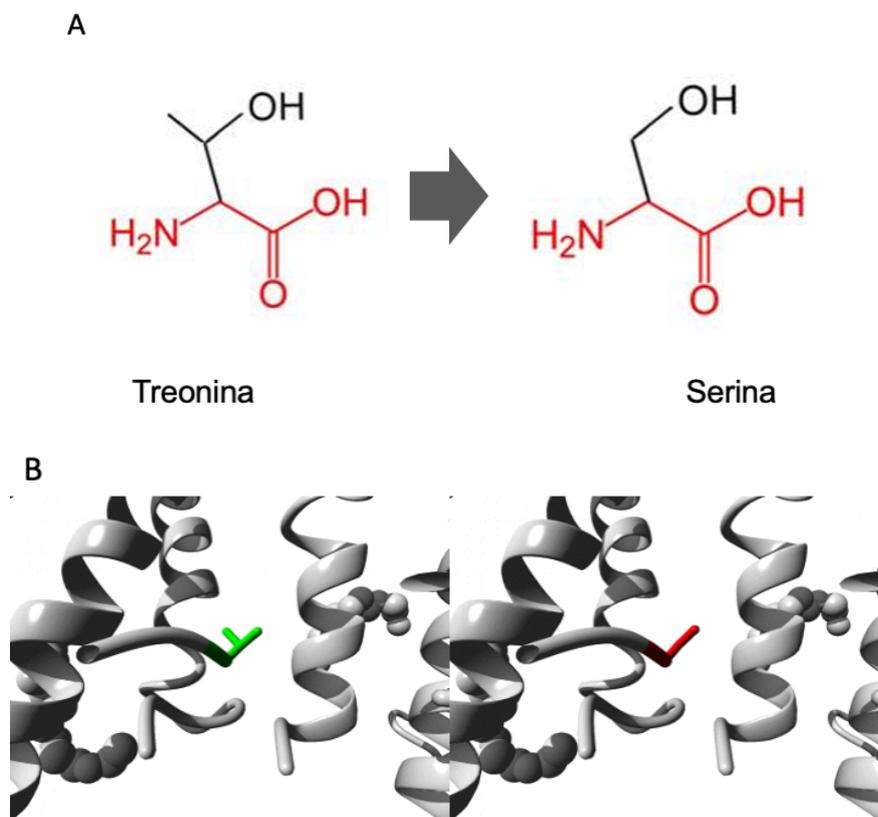


Figura 8: Imagens geradas pela ferramenta *HOPE* para a troca *TSHR_p.T574S*. A: Representação molecular dos aminoácidos treonina e serina, em vermelho a estrutura compartilhada, em preto a cadeia lateral; B: Modelo tridimensional de parte da proteína *TSHR* mostrando a diferença dos aminoácidos selvagem (verde) e mutado (vermelho).

FGF3_p.K101N

Identificamos uma única variante em homozigose no gene *FGF3*. A alteração *FGF3_p.K101N* identificada no paciente 20 não é descrita na literatura associada a fenótipos normais ou alterados, apenas dados sobre frequência foram encontrados na base do *dbSNP*.

A análise gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e afetando sítios de *splicing* que podem ter sua atividade aumentada, sem alterar as características da proteína. Observamos também que a região é bem conservada entre as espécies (Figura 9).

Gene FGF3 humano NM_005247	Proteína	p.K101N	
	cDNA	c.303G>C	
	gDNA	g.2684G>C	
Sítios de <i>splicing</i>			
Efeito	Posição	Pontuação	Borda exon-intron
Aceptor marginalmente aumentado	g.2679	wt: 0,2558 / mu: 0,2978	atga ACAA
Doador marginalmente aumentado	g.2685	wt: 0,8543 / mu: 0,9433	AAGA gggg
Doador aumentado	g.2686	wt: 0,89 / mu: 0,99	AGAG ggga
Conservação			
Espécies	Combinação	Alinhamento	
Humano	Referência	G L F S G R Y L A M N K R G R L Y A S E H Y S	
Alterado	Não conservada	G L F S G R Y L A M N N R G R L Y A S E H Y S	
<i>P. troglodytes</i>	Indêntica	G L F S G R Y L A M N K R G R L Y A S E H Y S	
<i>M. mulatta</i>	Indêntica	G L F S G R Y L A M N K R G R L Y A S E H Y S	
<i>F. catus</i>	Não homóloga		
<i>M. musculus</i>	Indêntica	G L F S G R Y L A M N K R G R L Y A S D H Y N	
<i>G. gallus</i>	Indêntica	G L F S G R Y L A M N K R G R L Y A S E N Y N	
<i>T. rubripes</i>	Não conservada	G L F S G R Y L A M N D K G R L Y A S E V F N	
<i>D. rerio</i>	Conservada	G L F S G R Y L A M N E K G R	
<i>D. melanogaster</i>	Não homóloga		
<i>C. elegans</i>	Não homóloga		
<i>X. tropicalis</i>	Conservada	G L F S G R Y L A M N Q R G R L Y A S E T Y N	

Figura 9: Quadro informativo com dados gerados pela ferramenta *MutationTaster* sobre a variante p.K101N no gene *FGF3*. wt: wild type – selvagem; mu: mutado; Pontuação dos sítios de *splicing*: valores maiores que 0,3 representam modificação dos sítios.

Análise gerada pela ferramenta *HOPE* prediz diversas consequências que podem ser causadas pela troca:

- **Aminoácidos.** Os aminoácidos de tipo selvagem e mutantes diferem em tamanho. O resíduo mutante é menor, isso pode levar a perda de interações. Há uma diferença de carga entre o aminoácido do selvagem e o mutante. A

carga de resíduo de tipo selvagem foi POSITIVA, a carga de resíduo mutante é NEUTRA. A carga do resíduo de tipo selvagem será perdida, o que pode causar perda de interações com outras moléculas ou resíduos (Figura 10).

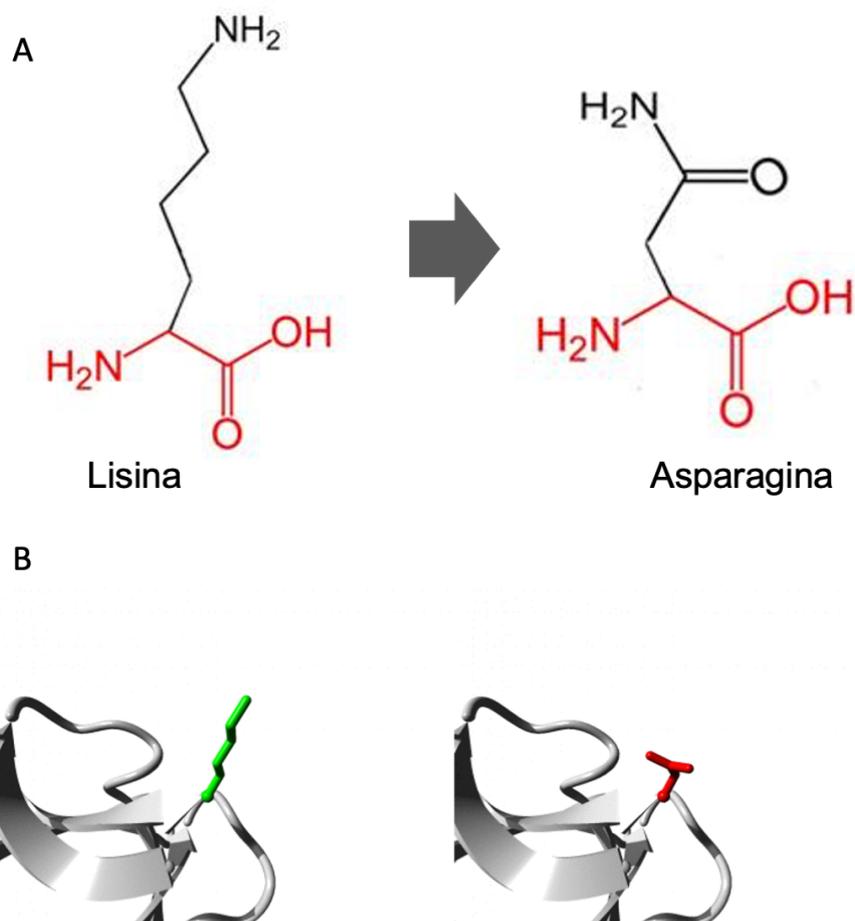


Figura 10: Imagens geradas pela ferramenta *HOPE* para a troca *FGF3_p.K101N*. A: Representação molecular dos aminoácidos lisina e asparagina, em vermelho a estrutura compartilhada, em preto a cadeia lateral; B: Modelo tridimensional de parte da proteína *FGF3* mostrando a diferença dos aminoácidos selvagem (verde) e mutado (vermelho).

- **Conservação.** O resíduo de tipo selvagem ocorre frequentemente nesta posição na sequência, mas outros resíduos também foram observados aqui. O resíduo mutante não estava entre os outros tipos de resíduos observados nesta posição em outras sequências homólogas. No entanto, resíduos com propriedades semelhantes foram observados, então é possível que seu resíduo mutante seja aceito aqui também. O resíduo mutante está localizado perto de uma posição altamente conservada.

- **Domínios.** O resíduo mutado está localizado em domínios importantes para a atividade principal da proteína. A mutação do resíduo pode perturbar essa função.

SFRP2_p.S235X

Também identificamos uma única variante com inserção de códon de parada no gene *SFRP2*. A alteração *SFRP2_p.S235X* identificada no paciente 35 em heterozigose não é descrita na literatura. Por ser uma alteração que causa a codificação de uma proteína truncada, já é espada a perda de função e deficiência da atividade de SFRP.

Gene SFRP2 humano NM_003013	Proteína	p.S235X
	cDNA	c.704C>A
	gDNA	g.7486C>A
Sítios de <i>splicing</i>		
Efeito	Posição	Pontuação
Acceptor marginalmente aumentado	g.7480	wt: 0,8644 / mu: 0,8797
Ganho de doador	g.7482	mu: 0,8800
Borda exon-intron		
		TGAA gaaa
		AAGA aata
Características da proteína		
Início (aa)	Final (aa)	Característica
172	295	Domínio NTR
245	245	Ponte dissulfeto
247	247	Ponte dissulfeto
289	289	Resíduo modificado, fosfoserina
295	295	Ponte dissulfeto
Efeito		
		Perda
Conservação		
Espécies	Combinação	Alinhamento
Humano	Referência	L N G V S E R D L K K S V L W L K D S L Q C T
Alterado	Não alinhada	
<i>P. troglodytes</i>	Conservada	L N G V S E R D L K K S V L W L K D S L Q C T
<i>M. mulatta</i>	Conservada	E R D L K K S V L W L K D S L Q C T
<i>F. catus</i>	Não homóloga	
<i>M. musculus</i>	Conservada	L N G V S E R D L K K S V L W L K D S L Q C T
<i>G. gallus</i>	Parc conservada	L N G L T E R D L R K I V L W L K G G L Q C T
<i>T. rubripes</i>	Parc conservada	L N G V T E R D L R K T V L W L K D G L Q C I
<i>D. rerio</i>	Parc conservada	T N G G V T D H D F R K T V L W L K D G L Q C V
<i>D. melanogaster</i>	Não homóloga	
<i>C. elegans</i>	Parc conservada	R R G V S D V E I R L S - - - S E S D G S C P
<i>X. tropicalis</i>	Parc conservada	L N G V T E R D L K K T V L W L K D G L Q C T

Figura 11: Quadro informativo com dados gerados pela ferramenta *MutationTaster* sobre a variante p.S235X no gene *SFRP2*. wt: wild type – selvagem; mu: mutado; Pontuação dos sítios de *splicing*: valores maiores que 0,3 representam modificação dos sítios.

A análise gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e afetando sítios de *splicing*, com perda de domínios e outras características importantes da proteína. A inserção prematura do códon de parada altera o tamanho da proteína, com perda de mais de 10% da sequência (de 888 para 705 aminoácidos). Observamos também que a região é parcialmente conservada entre as espécies (Figura 11).

As análises *in silico* mostram que as alterações *TSHR_p.T574S*, *FGF3_p.K101N* e *SFRP2_p.S235X* são possivelmente deletérias. Ensaio funcionais futuros poderão trazer maiores informações sobre as consequências destas variantes durante a embriogênese tireoidiana e seu envolvimento no fenótipo dos pacientes.

4.2.2 Genes regulados por fatores de transcrição

Após as análises de defeitos em genes candidatos, voltamos nossas buscas para outras regiões que poderiam estar envolvidas na causa da doença dos nossos pacientes. Os FTs *NKX2.1*, *FOXE1*, *PAX8* e *HHEX* estão presentes desde os primeiros estágios da embriogênese tireoidiana (7), contudo, nossa análise anterior não mostrou estes FTs afetados por variantes raras e deletérias. Direcionamos então nossas investigações a avaliar se a ação deles estaria comprometida. Selecionamos o *FOXE1* e o *PAX8* e analisamos se haveriam alterações nos genes que são controlados por estes FTs.

Compilamos listas de genes que são regulados pelos FTs *FOXE1* (Anexo 6) e *PAX8* (Anexo 7). Primeiramente, buscamos por variantes presentes na região promotora desses genes, local dos *motifs* de ligação dos FTs. Esta análise resultou em 63 variantes em genes regulados por *FOXE1* e 134 variantes em genes regulados por *PAX8*. Aplicamos filtros de frequência com dados populacionais dos bancos *ExAc*, *1000G* e *ABraOM*, selecionando as variantes raras (frequência menor que 1%). Verificamos a presença de 10 variantes raras em 8 genes regulados por *FOXE1* e 30 variantes raras em 17 genes regulados por *PAX8* (Tabela 5). Com as variantes raras selecionadas, utilizamos a ferramenta *TFBind* para verificar se as alterações estavam presentes dentro da região de ligação dos FTs, contudo, não identificamos variantes raras dentro de *motifs* de ligação dos FTs *FOXE1* e *PAX8*.

Tabela 5: Variantes raras na região 5'UTR de genes com *motif* para os fatores de transcrição FOXE1 e PAX8 identificadas nas amostras de disgenesias sequenciadas por WES

<i>MOTIF</i>	Pt	Gene_variante	<i>dbSNP</i>	<i>TFBind</i>
FOXE1	8	<i>FGF18_c.-207G>C</i>	-	N
FOXE1	10	<i>ANKRD37_c.-80C>A</i>	rs371823357	N
FOXE1	11	<i>FOLR1_c.-1596A>T</i>	rs187007568	N
FOXE1	14	<i>ETV5_c.-3085A>C</i>	-	N
FOXE1	22	<i>FOLR1_c.-1596A>T</i>	rs187007568	N
FOXE1	28	<i>SEC23B_c.-2718C>T</i>	rs749023698	N
FOXE1	29	<i>SEC23B_c.-3022C>T</i>	-	N
FOXE1	32	<i>DDIT3_c.-3011C>T</i>	rs28382838	N
FOXE1	32	<i>IGF2BP2_c.-22C>G</i>	rs373340007	N
FOXE1	35	<i>S100A4_c.-698C>T</i>	rs61805735	N
PAX8	1/2/12/27/34	<i>ANKRD9_c.-1753C>A</i>	rs184322569	N
PAX8	9	<i>FGFR2_c.-108C>T</i>	rs540720034	N
PAX8	10	<i>GALK2_c.-204C>T</i>	rs376158843	N
PAX8	12	<i>CITED2_c.-551G>A</i>	rs41289825	N
PAX8	15	<i>EIF4E_c.-312G>A</i>	rs186173714	N
PAX8	16/28	<i>OPRK1_c.-16562_-16571delinsC</i>	-	N
PAX8	17	<i>EIF4E_c.-664G>T</i>	-	N
PAX8	17	<i>FOXE1_c.-213C>T</i>	-	N
PAX8	18	<i>NUP107_c.-67C>T</i>	-	N
PAX8	19	<i>TRIB1_c.-1225A>G</i>	-	N
PAX8	20	<i>ARHGAP22_c.-221C>T</i>	rs548912881	N
PAX8	21	<i>SMIM22_c.-7015T>C</i>	rs79151406	N
PAX8	22	<i>TRIB1_c.-1188G>GA</i>	rs71295858	N
PAX8	26	<i>RASL10A_c.-466G>T</i>	-	N
PAX8	27	<i>GCSH_c.-87C>T</i>	rs372830976	N
PAX8	28	<i>FGFR2_c.-74G>A</i>	rs4647922	N
PAX8	29	<i>CRYAB_c.-154_-155delinsA</i>	-	N
PAX8	29	<i>PRR5L_c.-26G>A</i>	rs779506672	N
PAX8	30	<i>EIF4E_c.-94A>C</i>	rs368668353	N
PAX8	31	<i>ANKRD9_c.-1659T>A</i>	-	N
PAX8	32	<i>NR3C2_c.-5439C>G</i>	rs184722093	N
PAX8	32	<i>RASL10A_c.-112A>G</i>	rs561465325	N
PAX8	32	<i>RASL10A_c.-199G>C</i>	rs532086340	N
PAX8	33	<i>POMT1_c.-3218G>C</i>	rs556267289	N
PAX8	35	<i>EIF4E_c.-963_-964delinsG</i>	rs532888415	N

Pt: pacientes; N: neutro.

Em seguida, verificamos a presença de variantes na região codificadora dos genes regulados por PAX8 e FOXE1, buscando por alterações que prejudicariam a correta transcrição e tradução das proteínas. Esta análise resultou em 91 variantes em genes regulados por FOXE1 e 169 variantes em genes regulados por PAX8. Aplicamos filtros de frequência com dados populacionais dos bancos *ExAc*, *1000G* e *ABraOM*, selecionando as variantes raras (frequência menor que 1%). Com as ferramentas *FathMM-MKL*, *MutationTaster*, *Condel* e *PolyPhen2* anotamos a predição das variantes e selecionamos aquelas com potencial deletério. Esta linha de análise resultou em 14 variantes raras e deletérias em 10 genes regulados por FOXE1 e 29 variantes raras e deletérias em 18 genes regulados por PAX8 (Tabela 6).

Tabela 6: Variantes raras e deletérias na região codificadora de genes com *motif* para os fatores de transcrição FOXE1 e PAX8 identificadas no exoma das 35 amostras de DT sequenciadas (continua)

Pt	MOTIF	Gene_variante	Função	dbSNP	Predição			
					FAMM	COND	PPH2	MT
2	PAX8	<i>IGFBP7</i> _p.G124S	NS	rs371933194	D	N	D	D
	PAX8	<i>TG</i> _p.S1506N	NS	-	D	N	D	N
3	FOXE1	<i>NR4A2</i> _p.P146S	NS	rs374632208	D	N	N	D
4	PAX8	<i>FAM13A</i> _p.R50W	NS	rs114435452	D	N	D	N
5	FOXE1	<i>AMIGO3</i> _p.H361Y	NS	-	D	D	N	D
6	FOXE1	<i>DUOX2</i> _p.T423I	NS	rs201197899	D	D	N	D
	FOXE1	<i>DUSP5</i> _p.H224R	NS	rs113559164	D	N	N	D
8	PAX8	<i>LACTB</i> _p.I302V	NS	rs142381405	D	N	N	D
9	PAX8	<i>EGR1</i> _p.P218L	NS	-	D	N	D	D
10	FOXE1	<i>ADAMTS9</i> _p.P1924L	NS	-	D	N	D	D
	PAX8							
	PAX8	<i>POMT1</i>_p.S398R	NS	rs150367385	D	D	N	N
	PAX8	<i>TEKT4</i> _p.P244T ^a	NS	-	D	N	D	D
12	FOXE1	<i>DERL3</i> _c.314_315C	FS	rs536282991	-	-	-	D
	FOXE1	<i>DUOX2</i>_p.F966SfsX29	FS	rs530719719	-	-	-	D
14	PAX8	<i>TEKT4</i> _p.P244T ^a	NS	-	D	N	D	D
15	PAX8	<i>CAMK1G</i> _p.S10G	NS	rs138495087	D	N	N	D
16	FOXE1	<i>PRIMA1</i> _p.S91L	NS	-	D	D	D	D
18	PAX8	<i>ANKRD9</i> _p.P115L	NS	rs142725331	D	D	D	D
19	FOXE1	<i>SDF2L1</i> _p.A158T	NS	-	D	N	N	D

Tabela 6: Variantes raras e deletérias na região codificadora de genes com *motif* para os fatores de transcrição FOXE1 e PAX8 identificadas no exoma das 35 amostras de DT sequenciadas (conclusão)

Pt	MOTIF	Gene_variante	Função	dbSNP	Predição			
					FAMM	COND	PPH2	MT
	PAX8	LACTB_p.A338T	NS	rs186932521	D	D	N	D
	PAX8	LRRC58_p.E169G	NS	rs140152656	D	N	D	D
22	PAX8	NUP107_p.H757R	NS	rs145995997	D	D	D	D
	PAX8	TEKT4_p.S167L	NS	-	N	N	D	D
	PAX8	TEKT4_p.R169W	NS	-	N	D	D	D
24	FOXE1	AHCY_p.G123R	NS	rs41301825	D	N	D	D
26	FOXE1	BCAM_p.R227C	NS	rs150474390	N	D	D	N
27	FOXE1	BCAM_p.G575E	NS	-	D	N	D	D
28	PAX8	DGAT2_c.11delinsTCA	FS	-	-	-	-	D
29	PAX8	KRT7_p.R64Cb	NS	rs138391469	D	D	D	D
30	PAX8	POMT1_p.V256L	NS	rs200508760	D	D	N	D
	PAX8	TG_p.G1195R	NS	rs138469414	D	D	D	D
31	PAX8	SPARC_p.N69S	NS	rs13359508	D	D	N	D
32	FOXE1	DUOX2_p.R516H	NS	rs143818717	D	D	D	D
	PAX8	TEKT4_p.P244T ^a	NS	-	D	N	D	D
33	FOXE1	ADAMTS9_p.F786L	NS	rs145243531	D	N	N	D
	PAX8							
	PAX8	LRP8_p.T635S	NS	rs144428724	D	N	-	D
	PAX8	RUNX2_p.D320G	NS	-	D	D	-	D
	PAX8	KRT7_p.R64Cb	NS	rs138391469	D	D	D	D
34	PAX8	SLC26A7_p.V154M	NS	rs145779342	D	N	-	D
	PAX8	TG_p.Q870H	NS	rs2229843	D	D	-	N
35	PAX8	SLC26A7_p.L48F	NS	rs113085952	D	D	-	D

Pt: paciente; a e b: variantes que se repetem; NS: troca de único nucleotídeo não sinônima; FS: indel causadora de frameshift; FAMM: *Fathmm-MKL*; COND: *Condel*; PPH2: *PolyPhen2*; MT: *MutationTaster*; D: deletério; N: Neutro; em negrito as variantes com fenótipo associado.

Todas as 43 variantes foram identificadas em heterozigose e 2 variantes aparecem em comum entre 2 ou 3 pacientes (*KRT7_p.R64C* e *TEKT4_p.P244T*). Apenas no paciente 22 observamos 2 variantes no mesmo gene (*TEKT4*). As listas dos genes regulados por FOXE1 e PAX8 contendo variantes raras e deletérias foram submetidas na ferramenta de análise *Enrichr* para verificar a presença de vias enriquecidas ou mecanismo associados aos genes. O programa *Enrichr* não gerou termos associados aos genes com valores de $p < 0,05$ (a ferramenta utiliza o teste exato de Fischer como método estatístico, indicando a probabilidade de um gene

pertencer ao processo biológico ou via). Verificamos dados na literatura sobre cada uma das 43 variantes selecionadas, buscando informações sobre em qual quadro clínico foram descritas, se seriam polimorfismos ou mutações, ou se estariam envolvidas com o desenvolvimento. A variante *POMT1*_rs150367385 já foi descrita em pacientes com síndrome de Walker-Warburg e distrofia muscular (101-103), enquanto as variantes *DUOX2*_rs530719719 e *TG*_rs2229843 foram identificadas em pacientes com defeito da organificação do iodeto (95-97) e bócio (98), respectivamente.

Os resultados nos mostram que os *motifs* de ligação de FOXE1 e PAX8 não estão alterados em nossos pacientes, porém os genes controlados pelos FTs contêm variantes com predição deletéria que podem ser exploradas futuramente.

4.2.3 Análise do gene *FOXE1*

Seguidamente, decidimos continuar investigando o fator de transcrição FOXE1, já que um estudo *genome-wide* revelou quatro polimorfismos (rs7850258, rs965513, rs925489 e rs10759944) localizados perto do gene *FOXE1* associados a HC (104). O estudo demonstra que não apenas mutações, mas também SNPs, podem estar envolvidos na regulação complexa do desenvolvimento e função tireoidiana. Buscamos por todas as alterações em toda a região codificadora, 5'UTR e 3'UTR do gene *FOXE1*. Identificamos 26 variantes, dentre elas 12 SNVs na região promotora 5'UTR, 1 indel na região codificadora e 13 SNVs na região regulatória 3'UTR. Anotamos dados de frequência do alelo menor (MAF) de cada variante, tanto do sequenciamento do exoma, quando os das populações das bases de dados *ABraOM* e *1000G* (Tabela 7).

Tabela 7: Variantes identificadas na região de FOXE1 no exoma das 35 amostras de DT sequenciadas (continua)

Região	Variante	dbSNP	MAF		
			Exoma	<i>ABraOM</i>	<i>1000G</i>
5'UTR	c.-645C>A	rs143226042	0,04	0,03	0,03
5'UTR	c.-644A>G	rs13302470	0,61	0,64	0,69
5'UTR	c.-639T>A	-	0,01	-	-
5'UTR	c.-605G>C	-	0,01	-	-

Tabela 7: Variantes identificadas na região de FOXE1 no exoma das 35 amostras de DT sequenciadas (conclusão)

Região	Variante	dbSNP	MAF		
			Exoma	ABraOM	1000G
5'UTR	c.-537G>C	rs7849497	0,94	0,82	0,88
5'UTR	c.-283A>G	rs1867277	0,79	0,64	0,69
5'UTR	c.-248C>A	rs1867278	0,79	0,64	0,69
5'UTR	c.-213C>T	-	0,01	-	-
5'UTR	c.-177G>C	rs549601899	0,01	<0,00	<0,00
5'UTR	c.-156T>C	rs1867279	0,94	0,82	0,88
5'UTR	c.-131G>C	rs1867280	0,76	0,64	0,69
5'UTR	c.-115G>A	-	0,01	-	-
Exon	c.505_510del	rs71369530	0,83	0,72	-
3'UTR	c.*57T>C	rs7046645	0,94	0,82	8,88
3'UTR	c.*115A>C	rs7043516	0,16	0,18	0,18
3'UTR	c.*161G>T	rs1443434	0,79	0,64	0,69
3'UTR	c.*265T>C	rs1443435	0,79	0,64	0,69
3'UTR	c.*331A>G	rs41274258	0,04	0,02	0,01
3'UTR	c.*631T>C	rs41274260	0,06	0,03	0,03
3'UTR	c.*829A>G	rs756807798	0,01	-	<0,00
3'UTR	c.*967C>T	rs558753998	0,01	<0,00	<0,00
3'UTR	c.*1106C>A	-	0,01	-	-
3'UTR	c.*1174G>A	rs41274262	0,01	<0,00	<0,00
3'UTR	c.*1260C>T	rs76305281	0,07	0,06	0,07
3'UTR	c.*1342G>A	rs10984009	0,07	0,11	0,10
3'UTR	c.*1361G>T	rs375842088	0,01	-	-

Destacado em negrito: variantes selecionadas para análise.

Optamos por genotipar três polimorfismos: rs1867277 (presente em região promotora) e rs71369530 (inere no tamanho da cadeia polialanina de FOXE1), que nos dados do WES eram maiores comparados aos dados das populações do ABraOM e 1000G (ambas as variantes já foram associadas a doenças tireoidianas); e rs907577, localizado em região upstream à FOXE1, devido sua presença dentro de um bloco de risco para doenças tireoidianas e por estar em desequilíbrio de ligação com a variante rs1867277 em casos de câncer de tireoide. Realizamos o sequenciamento pontualmente por Sanger (Tabela 8) e analisamos estatisticamente a associação e o risco da presença do alelo menor e padrões de herança de cada variante com a DT dos pacientes (Tabela 9).

Tabela 8: Genótipos identificados nas disgenesias, controles normais e populações das bases de dados *ABraOM* e *1000G* para os polimorfismos rs907577, rs1867277 e rs71369530 do gene *FOXE1*

	Genótipo	Disgenesias	Controles	<i>ABraOM</i>	<i>1000G</i>
rs907577	T/T	25 (43,1%)	30 (46,2%)	-	1221 (48,8%)
	T/C	27 (46,6%)	24 (36,9%)	-	994 (39,7%)
	C/C	6 (10,3%)	11 (16,9%)	-	289 (11,5%)
	Total	58	65	-	2504
rs1867277	G/G	11 (18,3%)	29 (36,7%)	472 (40,3%)	1220 (48,7%)
	G/A	28 (46,7%)	37 (46,8%)	546 (46,6%)	995 (39,8%)
	A/A	21 (35,0%)	13 (16,5%)	153 (13,1%)	289 (11,5%)
	Total	60	79	1171	2504
rs71369530	16/16	2 (3,4%)	10 (9,2%)	82 (7,7%)	-
	16/14	17 (28,8%)	36 (33,0%)	425 (40,2%)	-
	14/14	40 (67,8%)	63 (57,8%)	551 (52,1%)	-
	Total	59	109	1058	-

Os valores se referem ao número de indivíduos genotipados para cada variante. Na base de dados *ABraOM* estão presentes os dados de genotipagem das variantes rs1867277 e rs71369530 e na base de dados *1000G* os das variantes rs907577 e rs1867277.

Na análise da variante rs907577, a frequência do alelo menor C nos pacientes com DT foi semelhante em comparação com a frequência no grupo de controles normais (CN) e no *1000G* (33,6%, 35,4% e 31,4%, respectivamente; $p > 0,05$) e não aumentou as chances para a doença (OR = 0,93, 95%CI = 0,55-1,57 e OR = 1,11, 95%CI = 0,75-1,64, respectivamente). O padrão de herança dominante (T/C + C/C) apareceu em 56,9% do grupo DT, em 53,9% do grupo CN ($p = 0,7341$; OR = 1,13, 95%CI = 0,56-2,31) e em 51,2% no *1000G* ($p = 0,3940$; OR = 1,26, 95%CI = 0,74-2,13). O padrão recessivo (C/C) foi observado em 10,3% do grupo DT, 16,9% do grupo CN ($p = 0,2913$; OR = 0,57, 95%CI = 0,19-1,64) e 11,5% do *1000G* ($p = 0,7777$; OR = 0,88, 95%CI = 0,38-2,08). Nenhuma análise adicional foi realizada porque a frequência alélica do polimorfismo rs907577 não estava contida no banco de dados *ABraOM*. Assim, A variante rs907577 não foi associada à presença de DT em nossos pacientes.

Tabela 9: Análise dos polimorfismos rs907577, rs1867277 e rs71369530 do gene *FOXE1*

Comparação		Padrão		
rs907577		Alelo menor C	T/C + C/C	C/C
DT	n=58	33,6%	56,9%	10,3%
	n=65	35,4%	53,9%	16,9%
CN	<i>P</i>	0,7715	0,7341	0,2913
	OR (95%CI)	0,93 (0,55-1,57)	1,13 (0,56-2,31)	0,57 (0,19-1,64)
	n=2504	31,4%	51,2%	11,5%
1000G	<i>P</i>	0,6089	0,394	0,7777
	OR (95%CI)	1,11 (0,75-1,64)	1,26 (0,74-2,13)	0,88 (0,38-2,08)
rs1867277		Alelo menor A	G/A + A/A	A/A
DT	n=60	58,3%	81,7%	35,0%
	n=79	39,9%	63,3%	16,5%
CN	<i>P</i>	0,0023	0,0178	0,0118
	OR (95%CI)	2,11 (1,30-3,42)	2,58 (1,16-5,74)	2,73 (1,23-6,07)
	n=1171	36,4%	57,1%	13,1%
ABraOM	<i>P</i>	<0,0001	0,0004	<0,0001
	OR (95%CI)	2,45 (1,69-3,55)	3,14 (1,62-6,11)	3,58 (2,05-6,25)
	n=2504	31,4%	51,3%	11,5%
1000G	<i>P</i>	<0,0001	<0,0001	<0,0001
	OR (95%CI)	3,06 (2,12-4,42)	4,23 (2,19-8,18)	4,13 (2,39-7,12)
rs71369530		Alelo menor 14 ^{Ala}	16/14 + 14/14	14/14
DT	n=59	82,2%	96,6%	67,8%
	n=109	74,3%	90,8%	57,8%
CN	<i>P</i>	0,1004	0,1646	0,204
	OR (95%CI)	1,60 (0,91-2,80)	2,88 (0,61-13,61)	1,54 (0,79-2,99)
	n=1058	72,2%	92,3%	52,1%
ABraOM	<i>P</i>	0,0172	0,2164	0,0186
	OR (95%CI)	1,78 (1,10-2,88)	2,39 (0,57-9,99)	1,94 (1,11-3,39)

DT: disgenesias tireoidianas; CN: controles normais; Destacado em cinza: comparações com valores de *p* significativos.

Já na análise da variante rs1867277, a frequência do alelo menor A foi maior nos casos de DT em comparação com o grupo CN (58,3% e 39,9%, respectivamente; *p* = 0,0023) e aumentou as chances da doença com OR = 2,11 (95%CI = 1,30-3,42). A comparação com dados populacionais confirmou esses resultados, a frequência do

alelo A foi de 36,4% no *ABraOM* ($p < 0,0001$) e 31,4% no *1000G* ($p < 0,0001$), com OR = 2,45 (95%CI = 1,69-3,55) e OR = 3,06 (95%CI = 2,12-4,42), respectivamente. O padrão dominante (G/A + A/A) aparece em 81,7% dos casos de DT e em 63,3% do grupo CN ($p = 0,01781$) com OR = 2,58 (95%CI = 1,16-5,79), em 57,1% do *ABraOM* ($p = 0,0004$) com OR = 3,14 (95%CI = 1,62-6,11) e em 51,3% do *1000G* ($p < 0,0001$) com OR = 4,23 (95%CI = 2,19-8,18). O padrão recessivo (A/A) foi observado em 35,0% dos casos de DT e em 16,5% do grupo CN ($p = 0,0118$) aumentando a chance da doença (OR = 2,73, 95%CI = 1,23-6,07), em 13,1% no *ABraOM* ($p < 0,0001$) com OR = 3,58 (95%CI = 2,05-6,25) e em 11,5% no *1000G* ($p < 0,0001$), com OR = 4,13 (95%CI = 2,39-7,12). Por tanto, o polimorfismo rs1867277 se mostrou associado à presença de DT, aumentando o risco para a doença.

Na análise da variante rs71369530 verificamos que a frequência do alelo menor 14^{Ala} nos casos de DT e no grupo CN foram ligeiramente próximos (82,2% e 74,3%, respectivamente; $p = 0,1004$) e não aumentaram as chances de doença (OR = 1,60, 95%CI = 0,91-2,80). Em contrapartida, na base de dados *ABraOM* a frequência foi de 72,2% ($p = 0,0172$) e aumentou o risco para a doença com OR = 1,78 (95%CI = 1,10-2,88). O padrão de herança dominante (16/14 + 14/14) foi observado em 96,6% dos casos de DT, em 90,8% do grupo CN ($p = 0,1646$; OR = 2,88, 95%CI = 0,61-13,61) e em 92,3% no *ABraOM* ($p = 0,2164$; OR = 2,39, 95%CI = 0,57-9,99). Porém, o padrão recessivo (genótipo 14/14) nos casos de DT (67,8%) associou-se à doença em comparação com *ABraOM* (52,1%; $p = 0,0186$) e aumentou a chance da doença com OR = 1,94 (95%CI = 1,11-3,39). Não foram observados resultados semelhantes em relação ao grupo CN, que apresentou frequência de 57,8% ($p = 0,2040$) e com OR = 1,54 (95%CI = 0,79-2,99). Nenhuma análise adicional foi realizada porque nenhum dado do polimorfismo rs71369530 estava contido no banco de dados *1000G*. Dessa forma, a variante rs71369530 também apresentou associação com a DT em comparação com a base de dados *ABraOM*, mas não com o grupo CN.

Como as variantes rs1867277 e rs71369530 demonstraram estar associadas à ocorrência de defeitos no desenvolvimento tireoidiano, realizamos uma análise conjunta com os pacientes e os indivíduos controle em que genotipamos ambos os polimorfismos. Comparamos todas as combinações de padrões de herança dominante e recessivo para as duas variantes e observamos que os genótipos combinados se mostraram associados à DT e conferiram um risco ainda maior para a

doença (Tabela 10). Primeiro, analisamos o padrão dominante combinado para ambos os polimorfismos, observado em 78,0% dos casos de DT e em 55,0% do grupo CN ($p = 0,0115$) com OR = 2,90 (95%CI = 1,25-6,72). Em seguida, avaliamos a combinação padrão rs1867277 dominante e rs71369530 recessivo, observado em 52,0% no grupo DT e 26,7% no grupo CN ($p = 0,0065$) com OR = 2,98 (95%CI = 1,34-6,61). Além disso, a combinação padrão rs1867277 recessivo e rs71369530 dominante foi observada em 32,0% no grupo DT e 10,0% no grupo CN ($p = 0,0041$) com OR = 4,24 (95%CI = 1,51-11,9). Finalmente, analisamos o padrão recessivo combinado para ambos os polimorfismos, observado em 28,0% dos casos de DT e em 5,0% do grupo CN ($p = 0,0009$) com OR = 7,39 (95%CI = 1,98-27,5).

Tabela 10: Análise conjunta dos polimorfismos rs1867277 e rs71369530 do gene *FOXE1*

Padrão			DT	CN	DT vs. CN	
			n=50	n=60		
rs1867277	+	rs71369530	%	%	p	OR (95%CI)
Dominante	+	Dominante	78,0	55,0	0,0115	2,90 (1,25-6,72)
Dominante	+	Recessivo	52,0	26,7	0,0065	2,98 (1,34-6,61)
Recessivo	+	Dominante	32,0	10,0	0,0041	4,24 (1,51-11,9)
Recessivo	+	Recessivo	28,0	5,0	0,0009	7,39 (1,98-27,5)

DT: disgenesias tireoidianas; CN: controles normais; n: número de indivíduos.

Baseados nestes resultados de que as variantes rs1867277 e rs71369530 mostraram associação e conferiram risco para o desenvolvimento de DT nas análises individuais e conjunta, investigamos a região cromossômica que envolve esses dois polimorfismos nos dados do exoma das 35 DT sequenciadas. Seleccionamos outras quatro variantes com MAFs mais elevados (em comparação com as populações *ABraOM* e *1000G*) dentro da região que inclui rs1867277 e rs71369530 (variantes destacadas na tabela 7). Comparamos os MAFs dessas 6 variantes (rs7849497, rs1867277, rs1867278, rs1867279, rs1867280 e rs71369530) nas 35 DT sequenciadas por WES e bancos de dados (Tabela 11). A análise individual mostrou que a presença do alelo menor de cada uma das 6 variantes demonstrou associação com TD quando comparada com a base de dados *ABraOM* (valores de $p < 0,0001$ a 0,0487) e aumentou o risco para TD, com OR de 1,78 a 3,68 (com resultados significativos nos valores de 95%CI). No entanto, não observamos associação

significativa e risco aumentado quando comparado ao banco de dados *1000G* (valores de $p > 0,0500$).

Tabela 11: Análise da frequência dos polimorfismos rs7849497, rs1867277, rs1867278, rs1867279, rs1867280 e rs71369530 do gene *FOXE1* no exoma das 35 amostras de DT sequenciadas

Variante	Alelo menor	WES	ABraOM	WES vs. ABraOM		1000G	WES vs. 1000G	
		%	%	<i>p</i>	OR (95%CI)	%	<i>p</i>	OR (95%CI)
rs7849497	C	94,29	81,77	0,0071	3,68 (1,33-10,2)	88,30	0,1206	2,19 (0,79-6,02)
rs1867277	A	78,57	63,62	0,0102	2,10 (1,18-3,73)	68,59	0,0736	1,68 (0,95-2,98)
rs1867278	A	78,57	63,62	0,0102	2,10 (1,18-3,73)	68,59	0,0736	1,68 (0,95-2,98)
rs1867279	C	94,29	81,77	0,0071	3,68 (1,33-10,2)	88,30	0,1206	2,19 (0,79-6,02)
rs1867280	C	75,71	63,62	0,0378	1,78 (1,03-3,10)	68,61	0,2030	1,43 (0,82-2,47)
rs71369530	14 ^{Ala}	82,86	72,16	0,0487	1,86 (0,99-3,50)	-	-	-

Por último, analisamos as variantes presentes na região 3'UTR a fim de verificar se os sítios de ligação de microRNAs (miRNA) foram alterados. Na base de dados *TargetScan* buscamos por locais alvo de miRNAs em toda a região 3'UTR do gene *FOXE1*. Verificamos a presença de 1446 regiões de ligação de miRNAs, que inclui 1442 sítios não conservados e 4 sítios bem conservados. Das 13 variantes identificadas em região 3'UTR nos dados de exoma das 35 DT sequenciadas, 11 delas estavam localizadas em 30 sítios pouco conservados (Tabela 12). Não identificamos variantes presentes em sítios conservado, o que indica que estes continuam íntegros e a região regulatória 3'UTR não foi afetada.

Tabela 12: Variantes na região 3'UTR de *FOXE1* no exoma das 35 disgenesias sequenciadas presentes em sítios de ligação de miRNAs (continua)

Variante	miRNA	Posição em 3'UTR
c.*57T>C	hsa-miR-10b-3p	55-61
	hsa-miR-4330	56-63
c.*115A>C	hsa-miR-423-3p	111-118
c.*161G>T	hsa-miR-181d-3p	158-164
c.*331A>G	hsa-miR-5004-3p	331-337

Tabela 12: Variantes na região 3'UTR de FOXE1 no exoma das 35 disgenesias sequenciadas presentes em sítios de ligação de miRNAs (conclusão)

Variante	miRNA	Posição em 3'UTR
c.*631T>C	hsa-miR-4499	627-633
	hsa-miR-873-3p	628-634
	hsa-miR-3175	631-637
c.*829A>G	hsa-miR-154-5p	823-829
	hsa-miR-6509-5p	825-831
c.*1106C>A	hsa-miR-4664-5p	1103-1109
	hsa-miR-342-5p	1103-1109
	hsa-miR-7110-5p	1104-1110
	hsa-miR-6842-5p	1104-1110
	hsa-miR-6752-5p	1104-1110
	hsa-miR-6732-5p	1105-1111
	hsa-miR-6798-5p	1106-1112
c.*1174G>A	hsa-miR-888-3p	1170-1176
c.*1260C>T	hsa-miR-5009-5p	1256-1263
	hsa-miR-8058	1256-1263
	hsa-miR-1269b	1257-1263
	hsa-miR-4529-3p	1257-1263
	hsa-miR-1269a	1257-1263
c.*1342G>A	hsa-miR-6877-3p	1338-1344
	hsa-miR-6819-3p	1338-1344
	hsa-miR-2114-3p	1339-1345
	hsa-miR-6823-3p	1339-1345
	hsa-miR-505-5p	1341-1347
c.*1361G>T	hsa-miR-4766-5p	1355-1362
	hsa-miR-3200-5p	1358-1364

Os resultados desta etapa mostram que polimorfismos presentes na região promotora e codificadora do gene *FOXE1* foram associados com o desenvolvimento de DT, aumentando o risco para a doença.

4.2.4 Análise da família DUOX

Ainda explorando as variantes presentes em genes candidatos, decidimos investigar os genes que participam da produção hormonal tireoidiana, e não apenas aqueles envolvidos na organogênese. Alguns trabalhos já descreveram mutações em

genes de disormonogênese em pacientes com DT (52, 53). Wang et al. mostrou as proteínas DUOX como genes patogênicos, sendo *DUOXA2* o gene variante mais comum em disgenesias e *DUOX2* o gene variante mais comum em casos de glândula *in situ* (105). Sendo assim, optamos por investigar as variantes presentes nos genes da família DUOX.

Buscamos por alterações em toda a região codificadora e regulatória dos genes da família DUOX (*DUOX1*, *DUOXA1*, *DUOX2* e *DUOXA2*). Especificamos a região cromossômica que compreende cromossomo 15:45384848-45457774 e identificamos 25 variantes. Anotamos dados de frequência populacionais dos bancos *ExAc*, *1000G* e *ABraOM*. Informações sobre as variantes contidas no banco de dados *ClinVar* também foram anotados. Com as ferramentas *FathMM-MKL*, *MutationTaster*, *Condel* e *PolyPhen2* anotamos a predição das variantes e selecionamos aquelas com potencial deletério. Esta análise resultou em 6 variantes deletérias no gene *DUOX2* e 2 variantes deletérias no gene *DUOXA1* (

Tabela 13).

Tabela 13: Variantes deletérias selecionadas na região dos genes da família DUOX identificadas no exoma das 35 amostras de DT sequenciadas

Pt	Variante	Função	dbSNP	ClinVar	Predição			
					FAMM	COND	PPH2	MT
18	<i>DUOX2</i> _c.3516-1G>C	SP	rs764517475	P	D	-	-	D
12	<i>DUOX2</i> _p.N1170T*	NS	rs200459845	-	D	D	N	D
15	<i>DUOX2</i> _p.A1123T*	NS	rs113632824	I	D	D	D	D
12	<i>DUOX2</i> _p.F966SfsX29*	FS	rs530719719	P	-	-	-	D
32	<i>DUOX2</i> _p.R516H*	NS	rs143818717	-	D	D	D	D
6	<i>DUOX2</i> _p.T423I*	NS	rs201197899	-	D	D	N	D
15	<i>DUOXA1</i> _p.W399X	SG	rs112166546	-	N	N	N	D
16/29	<i>DUOXA1</i> _p.T213M	NS	rs149960164	-	D	D	D	D

Pt: paciente; SP: troca em região de splicing; NS: troca de único nucleotídeo não sinônima; FS: indel causadora de frameshift; SG: stopgain; P: patogênico; FAMM: *Fathmm-MKL*; COND: *Condel*; PPH2: *PolyPhen2*; MT: *MutationTaster*; D: deletério; N: Neutro; * variantes confirmadas nos pacientes.

As 8 variantes foram identificadas em heterozigose nos pacientes. Com estes resultados, decidimos explorar mais as 6 variantes do gene *DUOX2*, com base em estudos que mostram variantes no gene *DUOX2* associadas à casos de disgenesia tireoidiana (53, 105). As 6 variantes foram pontualmente sequenciadas por Sanger e confirmamos 5 delas (Tabela 13) nos pacientes. As alterações se localizam em

diversos locais da proteína DUOX2 (Figura 12), exploramos possíveis efeitos deletérios destas variantes através de análises *in silico*.

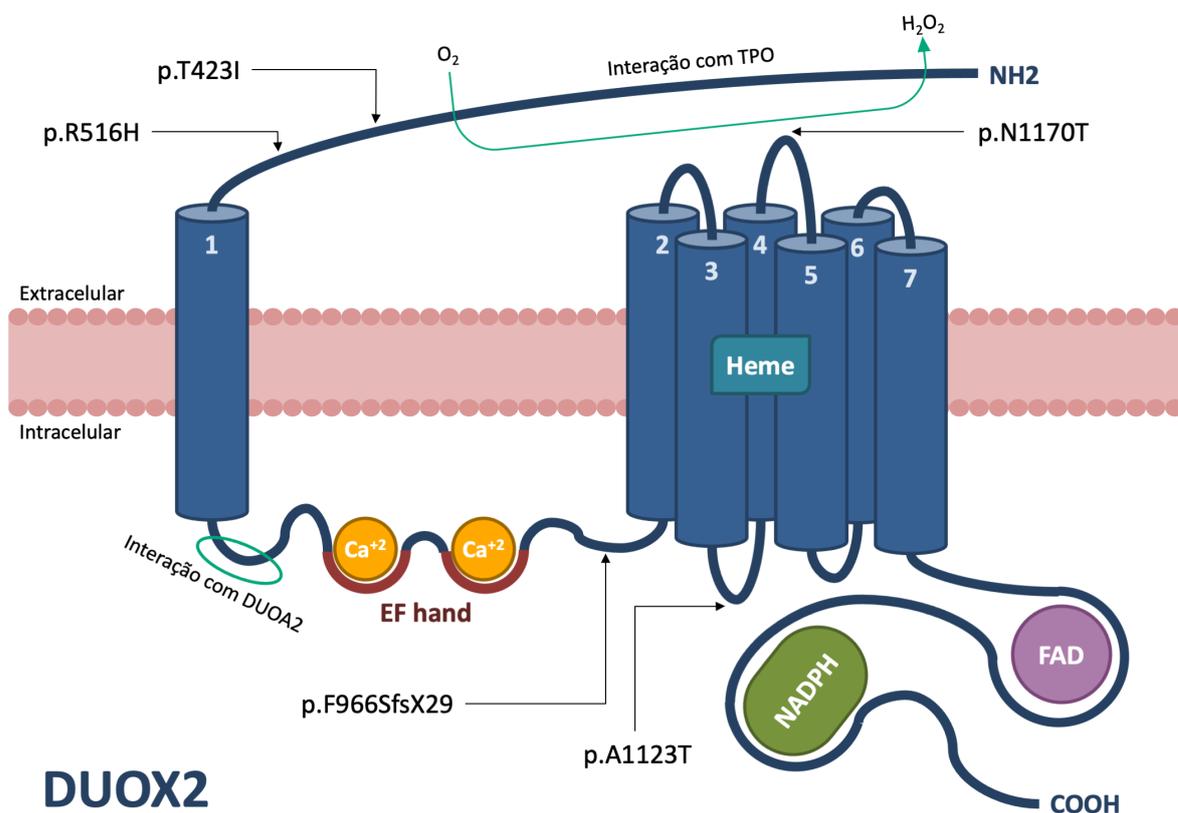


Figura 12: Representação esquemática da proteína DUOX2 na membrana apical do tireócito com a localização aproximada das variantes identificadas e confirmadas nas disgenesias tireoidianas.

DUOX2_p.T423I

A troca causada pela variante foi classificada com médio impacto funcional na análise da ferramenta *MutationAssessor*. A análise gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e mostra importantes domínios na proteína que podem ser perdidos devido a presença da variante, além de sítios de *splicing* que podem ser perdidos ou ganhos. Observamos também que a região é pouco conservada entre as espécies (Figura 13).

Gene DUOX2 humano NM_001363711	Proteína	p.T423I
	cDNA	c.1268C>T
	gDNA	g.5426C>T

Sítios de <i>splicing</i>			
Efeito	Posição	Pontuação	Borda exon-intron
Aceptor marginalmente aumentado	g.5426	wt: 0,6161 / mu: 0,6488	gtac AGAC
Aceptor aumentado	g.5435	wt: 0,2200 / mu: 0,4000	atgt GGCC
Aceptor aumentado	g.5428	wt: 0,4300 / mu: 0,5000	acag ACTA
Doador marginalmente aumentado	g.5422	wt: 0,9625 / mu: 0,9928	TCCC gtac

Características da proteína			
Início (aa)	Final (aa)	Característica	Efeito
26	601	Domínio topológico extracelular	Perda
30	596	Região de interação com peroxidase	Perda
455	455	Sítio de glicosilação	Pode ser perdido
537	537	Sítio de glicosilação	Pode ser perdido
602	622	Domínio transmembrana	Pode ser perdido
623	1041	Domínio topológico citoplasmático	Pode ser perdido
819	854	Domínio EF-hand 1	Pode ser perdido
832	843	Domínio de ligação Ca 1	Pode ser perdido
855	890	Domínio EF-hand 2	Pode ser perdido
868	879	Domínio de ligação Ca 2	Pode ser perdido
899	934	Domínio EF-hand 3	Pode ser perdido
960	1245	Região de interação com TXNDC11	Pode ser perdido
1042	1062	Domínio transmembrana	Pode ser perdido
1063	1076	Domínio topológico extracelular	Pode ser perdido
1077	1097	Domínio transmembrana	Pode ser perdido
1084	1266	Domínio oxiredutase férrico	Pode ser perdido
1098	1148	Domínio topológico citoplasmático	Pode ser perdido
1149	1169	Domínio transmembrana	Pode ser perdido
1170	1185	Domínio topológico extracelular	Pode ser perdido
1186	1206	Domínio transmembrana	Pode ser perdido
1207	1223	Domínio topológico citoplasmático	Pode ser perdido
1224	1244	Domínio transmembrana	Pode ser perdido
1245	1265	Domínio transmembrana	Pode ser perdido
1266	1548	Domínio topológico citoplasmático	Pode ser perdido
1267	1373	Domínio de ligação FAD tipo FR	Pode ser perdido

Conservação																									
Espécies	Combinação	Alinhamento																							
Humano	Referência	D	Y	W	P	G	P	G	K	F	S	R	T	D	Y	V	A	S	S	I	Q	R	G	R	
Alterado	Não conservada	D	Y	W	P	G	P	G	K	F	S	R	I	D	Y	V	A	S	S	I	Q	R	G	R	
<i>P. troglodytes</i>	Indêntica											S	R	T	D	Y	V	A	S	S	I	Q	R	G	R
<i>M. mulatta</i>	Não homóloga																								
<i>F. catus</i>	Indêntica											S	R	T	D	Y	V	A	S	S	I	Q	R	G	R
<i>M. musculus</i>	Indêntica											S	R	T	D	Y	V	A	S	S	I	Q	R	G	R
<i>G. gallus</i>	Não conservada	D	Y	W	Y	G	P	L	K	Y	S	R	A	D	Y	V	A	S	W	L	Q	R	G	R	
<i>T. rubripes</i>	Indêntica	D	F	M	Y	G	P	L	R	F	T	R	T	D	L	V	A	V	T	V	Q				
<i>D. rerio</i>	Conservada	D	Y	M	Y	G	P	L	R	F	S	R	S	D	A	V	A	L	T	I	Q	R	G	R	
<i>D. melanogaster</i>	Não conservada	D	K	L	F	G	P	M	E	F	T	R	R	D	L	G									
<i>C. elegans</i>	Não conservada											S	R	L	D	V	V	A	S	S	I	M	R	G	R
<i>X. tropicalis</i>	Indêntica	D	Y	C	Y	G	Q	I	K	Y	S	R	T	D	L	F	A	A	S	I	L	H	A	R	

Figura 13: Quadro informativo com dados gerados pela ferramenta *MutationTaster* sobre a variante p.T423I no gene *DUOX2*.

Análise gerada pela ferramenta *HOPE* prediz diversas consequências que podem ser causadas pela troca:

- **Aminoácidos.** Os aminoácidos de tipo selvagem e mutantes diferem em tamanho. O resíduo mutante é maior, isso pode levar a desordens na estrutura. A hidrofobicidade do resíduo de tipo selvagem e mutante difere. A mutação introduz um resíduo mais hidrofóbico nesta posição. Isso pode resultar na perda de ligações de hidrogênio e/ou perturbar o dobramento correto (Figura 14).
- **Contatos.** O resíduo de tipo selvagem é anotado no *UniProt* para estar envolvido em uma ponte de cisteína, que é importante para a estabilidade da proteína. Apenas cisteínas podem fazer esse tipo de ligação, a mutação causa perda dessa interação e terá um efeito severo na estrutura 3D da proteína. Junto com a perda da ligação da cisteína, as diferenças entre o antigo e o novo resíduo podem causar desestabilização da estrutura.
- **Estrutura.** A mutação está localizada dentro de um trecho de resíduos anotados no *UniProt* como uma região especial: *peroxidase-like*; medeia a atividade da peroxidase. As diferenças nas propriedades dos aminoácidos podem perturbar esta região e sua função.
- **Conservação.** O resíduo de tipo selvagem ocorre frequentemente nesta posição na sequência, mas outros resíduos também foram observados aqui. O resíduo mutante não estava entre os outros tipos de resíduos observados nesta posição em outras sequências homólogas. No entanto, resíduos com propriedades semelhantes foram observados, então é possível que o resíduo mutante seja aceito aqui também. O resíduo mutante está localizado perto de uma posição altamente conservada.
- **Domínios.** O resíduo mutado está localizado em domínios importantes para a atividade principal da proteína. A mutação do resíduo pode perturbar essa função.

variante, além de sítios de *splicing* que podem ser perdidos ou ganhos. Observamos também que a região é altamente conservada entre as espécies (Figura 15).

Análise gerada pela ferramenta *HOPE* prediz diversas consequências que podem ser causadas pela troca:

- **Aminoácidos.** Há uma diferença de carga entre o aminoácido do tipo selvagem e o mutante. A carga do resíduo de tipo selvagem será perdida, o que pode causar perda de interações com outras moléculas ou resíduos. Os aminoácidos de tipo selvagem e mutantes diferem em tamanho. O resíduo mutante é menor, o que pode levar à perda de interações (Figura 16).

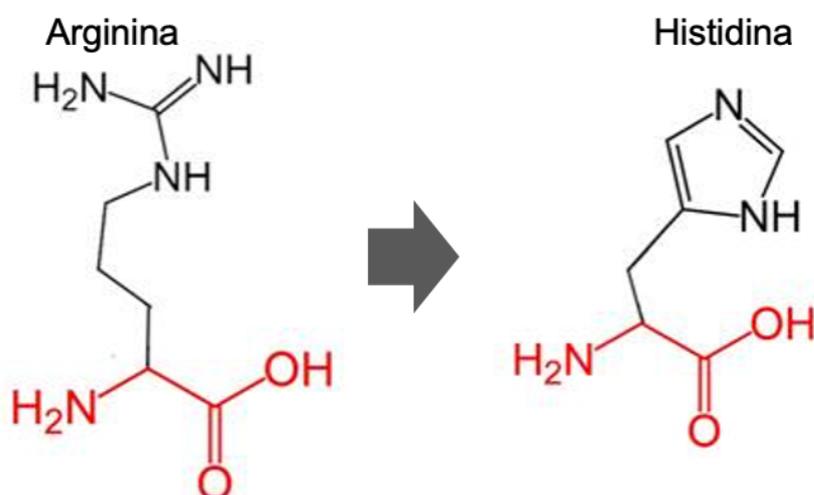


Figura 16: Representação molecular dos aminoácidos arginina e histidina na troca *DUOX2_p. p.R516H* (imagem gerada pela ferramenta *HOPE*). Em vermelho a estrutura compartilhada, em preto a cadeia lateral.

- **Contatos.** O resíduo de tipo selvagem é anotado no *UniProt* para estar envolvido em uma ponte de cisteína.
- **Estrutura.** A mutação está localizada dentro de um trecho de resíduos anotados no *UniProt* como uma região especial: peroxidase-like.
- **Conservação.** O resíduo de tipo selvagem é muito conservado, mas alguns outros tipos de resíduo também foram observados nesta posição. Nem o resíduo mutante nem outro tipo de resíduo com propriedades semelhantes foram observados nesta posição em outras sequências homólogas. Com base

nas pontuações de conservação, essa mutação provavelmente é prejudicial à proteína.

- **Domínios.** O resíduo mutado está localizado em um domínio importante para a atividade principal da proteína. A mutação do resíduo pode perturbar essa função.

DUOX2_p.F966SfsX29

A variante p.F966SfsX29 é descrita como causadora de defeitos na organificação do iodeto em pacientes com disormonogênese tireoidiana (95, 97). É uma deleção de quatro bases na sequência do gene DUOX2 (c.2895_2898del).

A análise gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e mostra importantes domínios na proteína que podem ser perdidos devido a presença da variante, além de sítios de *splicing* que podem ter suas atividades aumentadas. Observamos também que a região não é conservada entre as espécies. A deleção de bases causa alteração no quadro de aminoácidos e insere um códon de parada prematuro na posição 994, gerando uma proteína truncada de menor tamanho (de 1549 para 994 aminoácidos) (Figura 17).

Gene DUOX2 humano NM_001363711	Proteína	p.F966SfsX29		
	cDNA	c.2895_2898del		
	gDNA	g.13114_13117del		
Sítios de <i>splicing</i>				
Efeito	Posição	Pontuação	Borda exon-intron	
Aceptor marginalmente aumentado	g.13124	wt: 0,7246 / mu: 0,7624	actc GGAC	
Características da proteína				
Início (aa)	Final (aa)	Característica	Efeito	
623	1041	Domínio topológico citoplasmático	Perda	
960	1245	Região de interação com TXNDC11	Perda	
1042	1062	Domínio transmembrana	Perda	
1063	1076	Domínio topológico extracelular	Perda	
1077	1097	Domínio transmembrana	Perda	
1084	1266	Domínio oxiredutase férrico	Perda	
1098	1148	Domínio topológico citoplasmático	Perda	
1149	1169	Domínio transmembrana	Perda	
1170	1185	Domínio topológico extracelular	Perda	
1186	1206	Domínio transmembrana	Perda	
1207	1223	Domínio topológico citoplasmático	Perda	
1224	1244	Domínio transmembrana	Perda	
1245	1265	Domínio transmembrana	Perda	
1266	1548	Domínio topológico citoplasmático	Perda	
1267	1373	Domínio de ligação FAD tipo FR	Perda	
Conservação				
Espécie	Combinação	Alinhamento		
Human	Referência	F K Q N I S C R V S	F I T R T P G E R S H P Q	
Alterado	Não conservado	F K Q N I S C R V S	S L G H L - G S A P T P R	
<i>P. troglodytes</i>	Idêntica	S C R V S	F I T R T P G E R S H P Q	
<i>M. mulatta</i>	Não homólogo			
<i>F. catus</i>	Idêntica	S C R V S	F I T R T P G E R S C S Q	
<i>M. musculus</i>	Idêntica	A C R V S	F I N R T P G N R - - - -	
<i>G. gallus</i>	Idêntica	F K Q N V H N R V S	F I K K E P R G P S C K D	
<i>T. rubripes</i>	Idêntica	R V S	F I C - - P G S R F T T T	
<i>D. rerio</i>	Conservada	R K N S S S S V E	E L T H T P E E E H - - -	
<i>D. melanogaster</i>	Conservada	K Q N F L D T S T	N V A R M T S F N I E P -	
<i>C. elegans</i>	Idêntica	T Y D D F N A L F S	D I - - - - - - - - - -	
<i>X. tropicalis</i>	Conservada	T S A L I S	V L T K N T P Q N N H S N	

Figura 17: Quadro informativo com dados gerados pela ferramenta *MutationTaster* sobre a variante p.F966SfsX29 no gene *DUOX2*.

***DUOX2*_p.A1123T**

A troca causada pela variante foi classificada com médio impacto funcional na análise da ferramenta *MutationAssessor*. A análise gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e mostra

importantes domínios na proteína que podem ser perdidos devido a presença da variante, além de sítios de *splicing* que podem ser perdidos ou ganhos. Observamos também que a região é altamente conservada entre as espécies (Figura 18).

Gene DUOX2 humano NM_001363711	Proteína	p.A1123T	
	cDNA	c.3367G>A	
	gDNA	g.14635G>A	
Sítios de <i>splicing</i>			
Efeito	Posição	Pontuação	Borda exon-intron
Aceptor marginalmente aumentado	g.14629	wt: 0,4849 / mu: 0,5253	tttg ATGC
Doador aumentado	g.14637	wt: 0,7700 / mu: 0,9100	CGCA gtgg
Ganho de doador	g.14631	mu: 0,3400	TGAT gccca
Características da proteína			
Início (aa)	Final (aa)	Característica	Efeito
960	1245	Região de interação com TXNDC11	Perda
1084	1266	Domínio oxiredutase férrico	Perda
1098	1148	Domínio topológico citoplasmático	Perda
1149	1169	Domínio transmembrana	Pode ser perdido
1170	1185	Domínio topológico extracelular	Pode ser perdido
1186	1206	Domínio transmembrana	Pode ser perdido
1207	1223	Domínio topológico citoplasmático	Pode ser perdido
1224	1244	Domínio transmembrana	Pode ser perdido
1245	1265	Domínio transmembrana	Pode ser perdido
1266	1548	Domínio topológico citoplasmático	Pode ser perdido
1267	1373	Domínio de ligação FAD tipo FR	Pode ser perdido
Conservação			
Espécie	Combinação	Alinhamento	
Human	Referência	T F L N R Y V P F D A	A V D F H R W I A M A A
Alterado	Não conservado	T F L N R Y V P F D A	T V D F H R W I A M A A
<i>P. troglodytes</i>	Idêntica	T F L N R Y V P F D A	A V D F H R W I A M A A
<i>M. mulatta</i>	Não homólogo		
<i>F. catus</i>	Idêntica	T F L N R Y I P F D A	A V D F H R W I A M A A
<i>M. musculus</i>	Idêntica	T F L N R Y I P F D A	A V D F H R W I A M A A
<i>G. gallus</i>	Idêntica	T F L N R Y I P F D A	A V D F H R W I A M A A
<i>T. rubripes</i>	Idêntica	T F L N R Y I P F D A	A I D F H R F M A M A A
<i>D. rerio</i>	Idêntica	T F L N R Y I P F D A	A I D L H R Q M A A T A
<i>D. melanogaster</i>	Não conservado	F P I Q Q Y I P L D S	H I Q F H K I A A C T A
<i>C. elegans</i>	Idêntica		A I A F H K I V A L F A
<i>X. tropicalis</i>	Idêntica	T F L N Q Y I P F D S	A V D F H R L I A V T A

Figura 18: Quadro informativo com dados gerados pela ferramenta *MutaionTaster* sobre a variante p.A1123T no gene *DUOX2*.

Análise gerada pela ferramenta *HOPE* prediz diversas consequências que podem ser causadas pela troca:

- **Aminoácidos.** Os aminoácidos de tipo selvagem e mutantes diferem em tamanho. O resíduo mutante é maior, isso pode levar a desordens na estrutura. A hidrofobicidade do resíduo de tipo selvagem (apolar) e mutante (neutro) difere. As interações hidrofóbicas, seja no núcleo da proteína ou na superfície, podem ser perdidas (Figura 19).

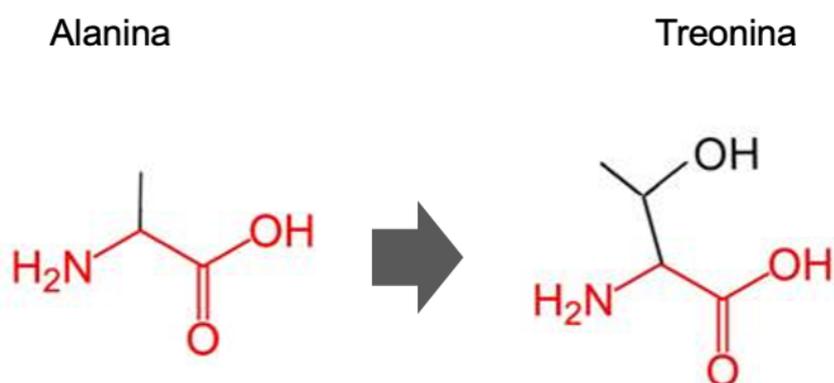


Figura 19: Representação molecular dos aminoácidos alanina e treonina na troca *DUOX2_p. A1123T* (imagem gerada pela ferramenta *HOPE*). Em vermelho a estrutura compartilhada, em preto a cadeia lateral.

- **Contatos.** O resíduo de tipo selvagem é anotado no *UniProt* para estar envolvido em uma ponte de cisteína, (semelhante a *DUOX2_p.T423I*).
- **Estrutura.** A mutação está localizada dentro de um domínio, anotado no *UniProt* como oxidoreductase férrica. A mutação introduz um aminoácido com propriedades diferentes, que pode perturbar esse domínio e abolir sua função. A mutação está localizada dentro de um trecho de resíduos anotados no *UniProt* como uma região especial: Interação com TXNDC11. As diferenças nas propriedades dos aminoácidos podem perturbar esta região e perturbar sua função.
- **Conservação.** O resíduo de tipo selvagem é muito conservado, mas alguns outros tipos de resíduo também foram observados nesta posição. Nem o resíduo mutante nem outro tipo de resíduo com propriedades semelhantes foram observados nesta posição em outras sequências homólogas. Com base nas pontuações de conservação, essa mutação provavelmente é prejudicial à proteína.

- **Domínios.** O resíduo mutado está localizado em um domínio importante para a atividade principal da proteína. A mutação do resíduo pode perturbar essa função.

DUOX2_p.N1170T

A troca causada pela variante foi classificada com baixo impacto funcional na análise da ferramenta *MutationAssessor*. A análise gerada pela ferramenta *MutationTaster* classifica a alteração como causadora de doença e mostra importantes sítios na proteína que podem ser perdidos devido a presença da variante, além de sítios de *splicing* que podem ser perdidos ou ganhos. Observamos também que a região é pouco conservada entre as espécies (Figura 20).

Análise gerada pela ferramenta *HOPE* prediz diversas consequências que podem ser causadas pela troca:

- **Aminoácidos.** Os aminoácidos de tipo selvagem e mutantes diferem em tamanho. O resíduo mutante é menor, o que pode levar à perda de interações. A hidrofobicidade do resíduo de tipo selvagem e mutante difere. A mutação introduz um resíduo mais hidrofóbico nesta posição. Isso pode resultar na perda de ligações de hidrogênio e/ou perturbar o dobramento correto (Figura 21).
- **Estrutura.** A mutação está localizada dentro de um domínio oxidoreductase férrico e de um trecho de resíduos de interação com TXNDC11.
- **Conservação.** O resíduo de tipo selvagem não é conservado nesta posição. O resíduo mutante está entre os tipos de resíduos observados nesta posição em outras sequências homólogas.
- **Domínios.** O resíduo mutado está localizado em um domínio importante para a atividade principal da proteína. A mutação do resíduo pode perturbar essa função.

Gene DUOX2 humano NM_001363711	Proteína	p.N1170T		
	cDNA	c.3509A>C		
	gDNA	g.14956A>C		
Sítios de splicing				
Efeito	Posição	Pontuação	Borda exon-intron	
Aceptor marginalmente aumentado	g.14953	wt: 0,8542 / mu: 0,8882	ttgt GAAT	
Doador marginalmente aumentado	g.14961	wt: 0,4734 / mu: 0,5451	GATG ggtc	
Doador aumentado	g.14955	wt: 0,3300 / mu: 0,6700	GTGA atga	
Ganho de acceptor	g.14958	mu: 0,3000	actg ATGG	
Características da proteína				
Início (aa)	Final (aa)	Característica	Efeito	
960	1245	Região de interação com TXNDC11	Perda	
1084	1266	Domínio oxiredutase férrico	Perda	
1170	1185	Domínio topológico extracelular	Perda	
1186	1206	Domínio transmembrana	Pode ser perdido	
1207	1223	Domínio topológico citoplasmático	Pode ser perdido	
1224	1244	Domínio transmembrana	Pode ser perdido	
1245	1265	Domínio transmembrana	Pode ser perdido	
1266	1548	Domínio topológico citoplasmático	Pode ser perdido	
1267	1373	Domínio de ligação FAD tipo FR	Pode ser perdido	
Conservação				
Espécies	Combinação	Alinhamento		
Humano	Referência	L L A C I F P N V F V	N	D G S K L P Q K F Y W
Alterado	Não conservada	L L A C I F P N V F V	T	D G S K L P Q K F Y W
<i>P. troglodytes</i>	Indêntica	L L A C V F P N V F V	N	D G S K L P Q K F Y W
<i>M. mulatta</i>	Não homóloga			
<i>F. catus</i>	Indêntica	L L A C V F P N V F V	N	D G S Q L P Q K F Y W
<i>M. musculus</i>	Indêntica	L M A C V F P N V F V	N	D G S K F P P K Y Y W
<i>G. gallus</i>	Indêntica	V L S C L F S S V F M	N	D G S Q L P Q K Y Y W
<i>T. rubripes</i>	Indêntica	I L S C L F P K V F R	N	N G S E R P M K W S W
<i>D. rerio</i>	Indêntica			P K V F S N N G S E L P M K W T F
<i>D. melanogaster</i>	Conservada	N L R C L T R E - - V	H	F A S D Y K P D I T F
<i>C. elegans</i>	Não conservada	G L A C L F Q E A F F	-	- G S N F L P S I S Y
<i>X. tropicalis</i>	Conservada	V L S C L F P T V F V	D	D G S D H P N K Y Y W

Figura 20: Quadro informativo com dados gerados pela ferramenta *MutationTaster* sobre a variante p.N1170T no gene *DUOX2*.

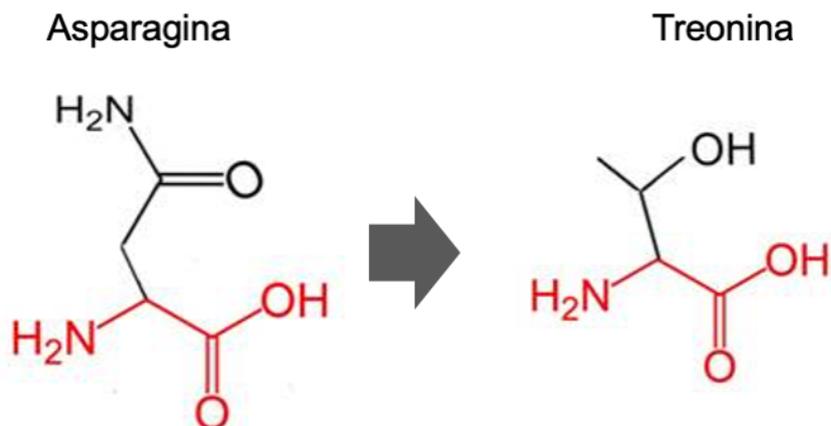


Figura 21: Representação molecular dos aminoácidos asparagina e treonina na troca DUOX2_p.N1170T (imagem gerada pela ferramenta HOPE). Em vermelho a estrutura compartilhada, em preto a cadeia lateral.

Os resultados desta etapa do estudo permitiram identificar variantes no gene *DUOX2* em pacientes com DT que possivelmente são deletérias podem afetar a ação da proteína DUOX2. Um estudo funcional está em andamento para fornecer mais informações sobre a ação dessas alterações na função da proteína.

4.3 Regiões em homozigose

Investigamos os dados de sequenciamento WES em busca de regiões em homozigose. Utilizamos os dados completo de todas as variantes identificadas nos pacientes, sem excluir regiões ou tipos de variantes para esta análise. Identificamos duas regiões em homozigose em dois pacientes (Tabela 14).

No paciente 14 identificamos uma região em homozigose no cromossomo 2, contendo 4 genes. No paciente 24 identificamos uma região em homozigose no cromossomo 20, contendo 5 genes. Analisamos as variantes contidas dentro dessas duas regiões em homozigose. Investigamos aquelas que poderiam ser funcionais dentro de regiões codificadoras, intrônicas adjacentes, 5'UTR e 3'UTR. Verificamos que nenhuma das variantes contidas nestes 9 genes presentes nas regiões em homozigose eram raras, todas apresentavam frequência maior que 1%. Portanto, a presença de homoziguidade não foi associada às DTs neste estudo.

Tabela 14: Variantes contidas nas regiões em homozigose identificadas no exoma das amostras sequenciadas dos pacientes 14 e 24 sequenciados por WES

	Posição Cromossômica	Região	Gene	Variante	dbSNP	MAF 1000G
Paciente 14	Chr2:234130	Intron	<i>SH3YL1</i>	.	rs3748872	0,45
	Chr2:239736	Intron	<i>SH3YL1</i>	.	rs10181015	0,49
	Chr2:240054	Intron	<i>SH3YL1</i>	.	rs2864044	0,50
	Chr2:240255	Intron	<i>SH3YL1</i>	.	rs300704	0,17
	Chr2:240630	Intron	<i>SH3YL1</i>	.	rs2085852	0,45
	Chr2:241669	Intron	<i>SH3YL1</i>	.	-	-
	Chr2:242800	Exon	<i>SH3YL1</i>	p.M1V	rs10181051	0,31
	Chr2:243504	5'UTR	<i>SH3YL1</i>	c.-704G>A	rs7556816	0,31
	Chr2:256278	5'UTR	<i>SH3YL1</i>	c.-13478A>G	rs4455191	0,28
	Chr2:260594	Intron	<i>SH3YL1</i>	.	rs4854299	0,01
	Chr2:260677	Intron	<i>SH3YL1</i>	.	rs10188763	0,45
	Chr2:262553	Intron	<i>SH3YL1</i>	.	rs7596678	0,50
	Chr2:263270	Intron	<i>SH3YL1</i>	.	rs10167992	0,16
	Chr2:264895	5'UTR	<i>ACP1</i>	c.-70G>C	rs4854300	0,01
	Chr2:269352	Intron	<i>ACP1</i>	.	rs6755722	0,45
	Chr2:272840	Intron	<i>ACP1</i>	.	rs12714401	0,50
	Chr2:272926	Intron	<i>ACP1</i>	.	rs12714402	0,50
	Chr2:273070	Intron	<i>ACP1</i>	.	rs12714403	0,17
	Chr2:279705	3'UTR	<i>FAM150B</i>	c.*442G>A	rs3828329	0,45
	Chr2:286812	Intron	<i>FAM150B</i>	.	rs4484059	0,17
Chr2:676177	Intron	<i>TMEM18</i>	.	rs2293083	0,50	
Paciente 24	Chr20:76962	Exon	<i>DEFB125</i>	p.S125S	rs6111385	0,23
	Chr20:138125	5'UTR	<i>DEFB127</i>	c.-61G>T	rs2298108	0,33
	Chr20:138148	5'UTR	<i>DEFB127</i>	c.-38C>T	rs2298109	0,33
	Chr20:139409	Intron	<i>DEFB127</i>	.	rs12624951	0,32
	Chr20:139456	Exon	<i>DEFB127</i>	p.G31R	rs12624954	0,32
	Chr20:139576	Exon	<i>DEFB127</i>	p.R71S	rs16995685	0,32
	Chr20:139745	3'UTR	<i>DEFB127</i>	c.*80T>C	rs4142304	0,07
	Chr20:207889	Upstream	<i>DEFB129</i>	.	rs2298149	0,27
	Chr20:210061	Exon	<i>DEFB129</i>	p.L67L	rs13045643	0,27
	Chr20:277692	Upstream	<i>ZCCHC3</i>	.	rs3746790	0,49
	Chr20:277832	Upstream	<i>ZCCHC3</i>	.	rs1997793	0,35
	Chr20:278515	Exon	<i>ZCCHC3</i>	p.D96D	rs2223665	0,41
	Chr20:309692	3'UTR	<i>SOX12</i>	c.*2176A>G	rs6050276	0,12

Com estes resultados, descartamos a hipótese de que estas regiões em homozigose poderiam estar associadas à causa da doença nos pacientes por não haver variantes raras nas regiões. Dessa forma, a presença de homoziguidade não foi associada às DTs neste estudo.

4.4 Variantes estruturais CNVs

Com a análise realizada com o programa XHMM, foram selecionadas 10 CNVs presentes em 12 pacientes. Foi feita uma análise detalhada de casa uma delas na base de dados *DGV*, e verificamos que nenhuma delas é descrita (Tabela 15).

Com as informações do *DGV* também observamos que nas mesmas regiões cromossômicas existem outras CNVs de inserção ou duplicação de maior tamanho. Somente duas CNVs (ambas duplicações) envolviam a sequência codificante de genes (genes *LOC283177* e *AHNAK2*). As outras CNVs localizavam-se em regiões de introns, não continham genes ou se encontravam fora da região codificadora de genes. A análise feita na base de dados *ClinVar* permitiu verificar que nenhuma CNV já descrita contendo exclusivamente os genes *LOC283177* e *AHNAK2* foi associada a doença ou considerada patogênica. Dessa forma, a presença de CNVs não foi associada às DTs neste estudo.

Tabela 15: Alterações estruturais (CNVs) identificadas no exoma das amostras de DT sequenciadas

Pt	CNV	Localização cromossômica	Gene na região
3	DEL	chr7:72418915-72430805	NSUN5P2
6	DEL	chr5:34193259-34193718	-
13	DUP	chr14:105417588-105418355	AHNAK2
15	DEL	chr14:105412780-105414503	AHNAK2
16	DUP	chr17:20355166-20361681	LGALS9B
16	DEL	chr5:34193259-34193718	-
20	DUP	chr1:155184142-155203274	GBAP1
20	DEL	chr7:100634878-100636333	-
29	DUP	chr7:100634878-100636333	-
31	DUP	chr11:134356668-134357175	LOC283177
33	DUP	chr5:34193259-34193718	-
35	DUP	chr8:11986329-12003959	FAM66D

Pt: paciente; DEL: deleção; DUP: duplicação. Em destaque as duas CNVs localizadas na região codificadora de genes.

4.5 Análise individual dos pacientes

Com os resultados das análises anteriores observamos variantes relevantes presentes no exoma de todas as 35 amostras disgenesia sequenciadas. Por fim, realizamos uma análise individualizada observando quais dessas alterações foram identificadas em cada paciente e comparamos os achados com o fenótipo do paciente e dados já descritos na literatura.

Utilizamos as listas das variantes resultantes das análises de “Genes Candidatos”, “Genes Regulados por FTs”, “Análise do Gene FOXE1” e “Análise da Família DUOX”. Seleccionamos as variantes presentes em genes diretamente relacionados com função, desenvolvimento e doenças tireoidianas, e também os genes envolvidos com desenvolvimento embrionário de outros tecidos (

Figura 22). Observamos que apenas no paciente 10 não identificamos variantes em genes diretamente relacionados com o desenvolvimento e/ou função tireoidiana. No entanto, o paciente apresenta variantes em genes que participam do desenvolvimento de outros tecidos e ainda genes que já foram descritos na etiologia do câncer de tireoide.

Utilizamos a ferramenta *String* e verificamos a interação dos genes alterados contendo as variantes selecionadas por paciente. Observamos interações nos genes alterados dos pacientes 1, 2, 3, 6 (Figura 23), 12, 14, 15, 18, 20 (Figura 24), 30, 31, 32, 33, 34 e 35 (Figura 25). A maioria das interações era de média confiança, sendo que as interações entre os genes *FOXE1_TPO*, *FOXE1_TG*, *FOXE1_TSHR*, *DUOX2_DUOXA1*, *SPARC_COL1A1*, *BMP4_JAG1* e *JAG1_RUNX2* apresentaram pontuação de alta confiança.

Estes resultados mostram que a maioria dos pacientes possui variantes em genes diretamente relacionados com o desenvolvimento e função tireoidiana. Embora sejam variantes heterozigotas e não descritas anteriormente em casos de defeito de desenvolvimento da tireoide, as análises de interação podem sugerir uma possível causa poligênica da DT.

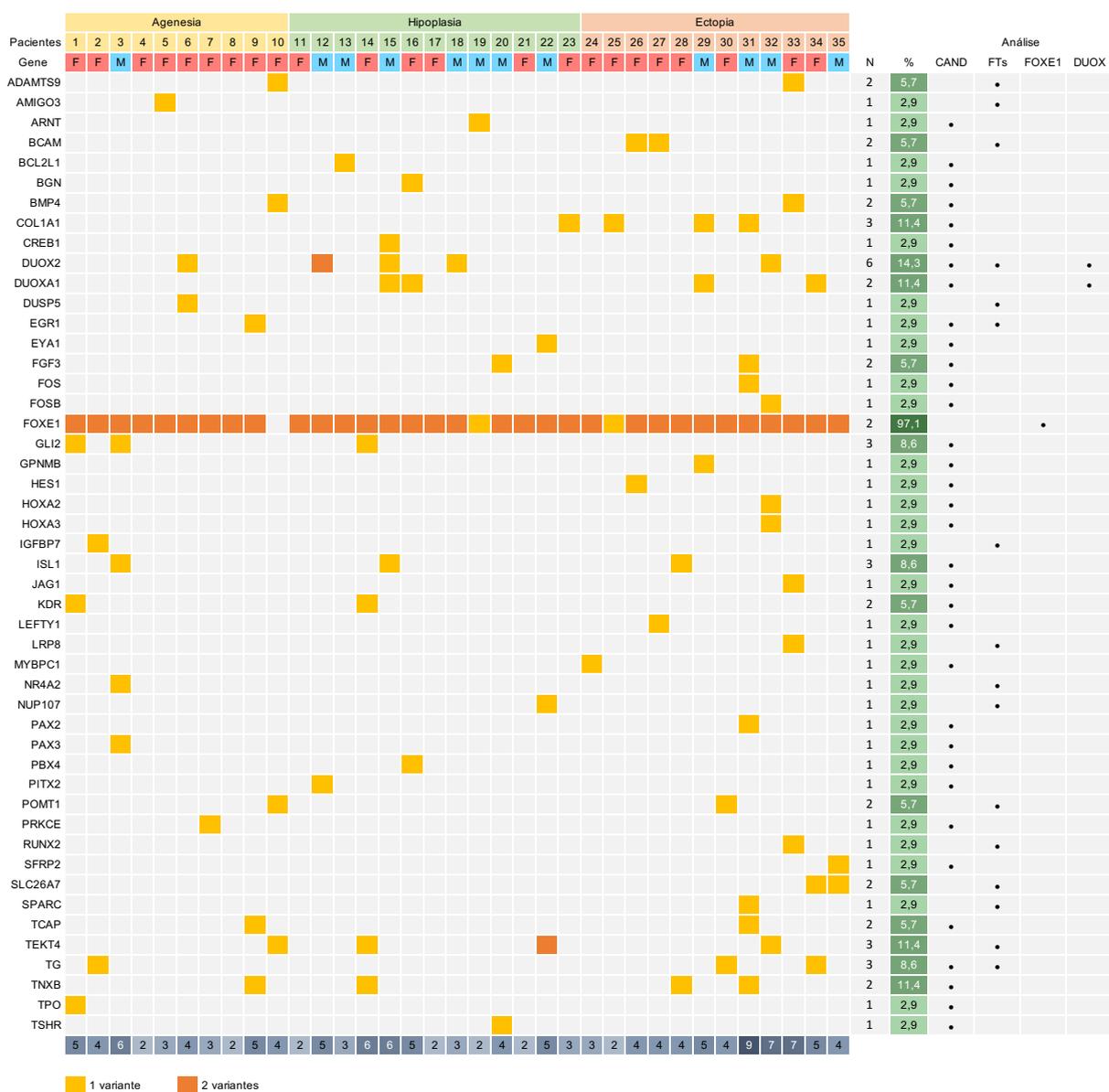


Figura 22: Distribuição das principais variantes identificadas no exoma das 35 amostras de DT sequenciadas. As cores nos genes indicam em qual análise foram identificadas as variantes. CAND: candidatos; FTs: fatores de transcrição; FOXE1: gene *FOXE1*; DUOX: genes da família DUOX; TOTAL: total de variantes por paciente; N: número de variantes identificadas no gene; %: frequência do gene alterado entre os pacientes.

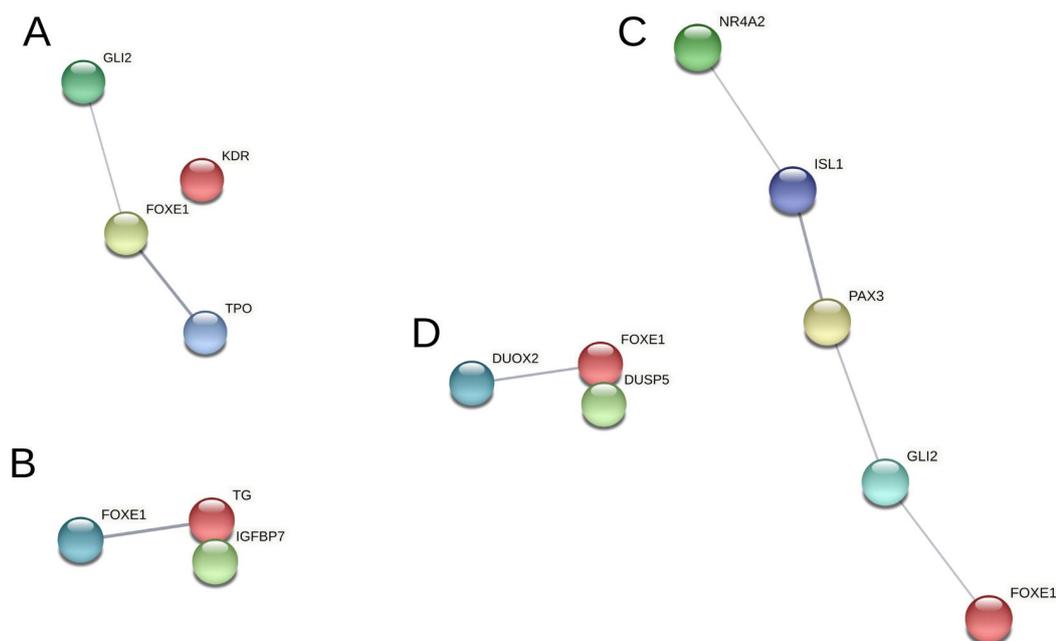


Figura 23: Interação dos genes alterados (análise por paciente) gerada pela ferramenta *String*. Pacientes com agenesia: A. Paciente 1, pontuação das interações (PI): *GLI2_FOXE1* = 0,440 e *FOXE1_TPO3* = 0,669; B. Paciente 2, PI: *FOXE1_TG* = 0,701; C. Paciente 3, PI: *FOXE1_GLI2* = 0,440, *GLI2_PAX3* = 0,461, *PAX3_ISL1* = 0,639 e *ISL1_NR4A2* = 0,422; D. Paciente 6, PI: *FOXE1_DUOX2* = 0,562. Pontuação das interações representam o grau de confiança: baixa (0,0 – 0,3), média (0,4 – 0,6) ou alta (0,7 – 1,0).

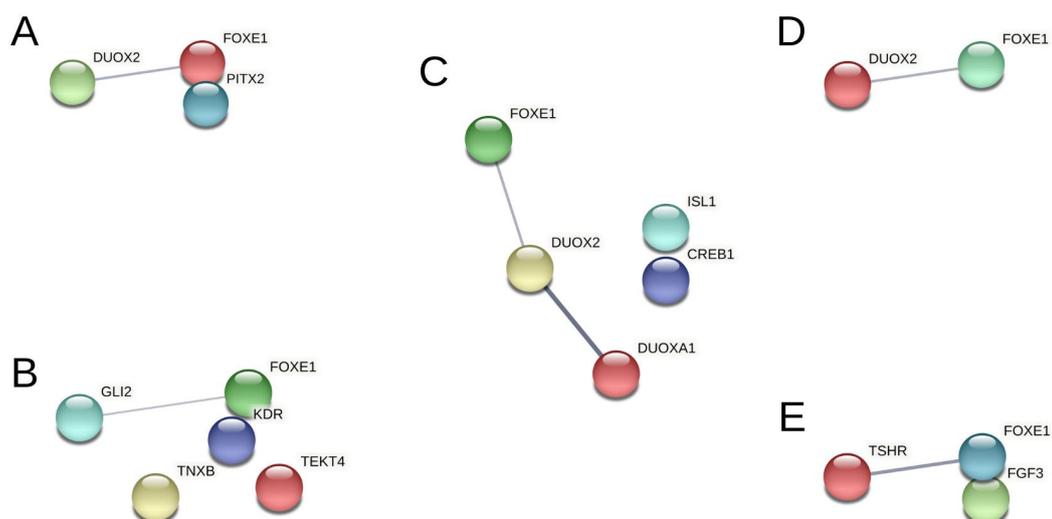


Figura 24: Interação dos genes alterados (análise por paciente) gerada pela ferramenta *String*. Pacientes com hipoplasia: A. Paciente 12, pontuação das interações (PI): *FOXE1_DUOX2* = 0,562; B. Paciente 14, PI: *GLI2_FOXE1* = 0,440; C. Paciente 15, PI: *FOXE1_DUOX2* = 0,562 e *DUOX2_DUOX1* = 0,980; D. Paciente 18, PI: *FOXE1_DUOX2* = 0,562; E. Paciente 20, PI: *FOXE1_TSHR* = 0,736. Pontuação das interações representam o grau de confiança: baixa (0,0 – 0,3), média (0,4 – 0,6) ou alta (0,7 – 1,0).

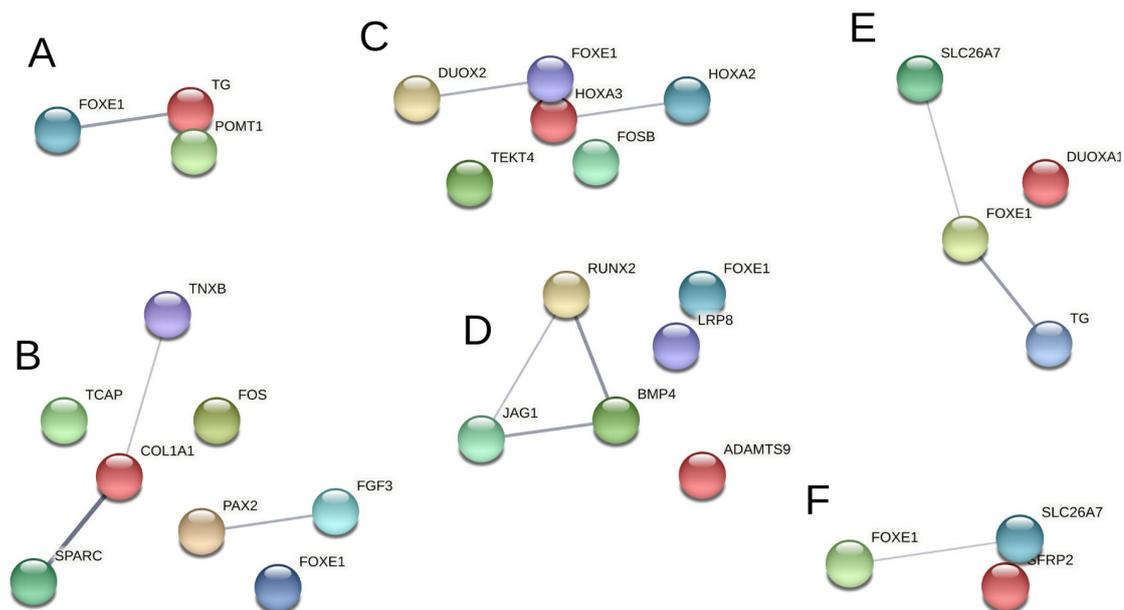


Figura 25: Interação dos genes alterados (análise por paciente) gerada pela ferramenta String. Pacientes com ectopia: A. Paciente 30, pontuação das interações (PI): $FOXE1_TG = 0,701$; B. Paciente 31, PI: $SPARC_COL1A1 = 0,989$, $COL1A1_TNXB = 0,433$ e $PAX2_FGF3 = 0,565$; C. Paciente 32, PI: $FOXE1_DUOX2 = 0,562$ e $HOXA3_HOXA2 = 0,531$; D. Paciente 33, PI: $BMP4_JAG1 = 0,664$, $JAG1_RUNX2 = 0,753$ e $RUNX2_BMP4 = 0,467$; E. Paciente 34, PI: $SLC26A7_FOXE1 = 0,415$ e $FOXE1_TG = 0,701$; F. Paciente 35, PI: $FOXE1_SLC26A7 = 0,415$. Pontuação das interações representam o grau de confiança: baixa (0,0 – 0,3), média (0,4 – 0,6) ou alta (0,7 – 1,0).

5 DISCUSSÃO

Neste estudo pesquisamos alterações gênicas com relevância clínica em pacientes com DT, cuidadosamente caracterizados em termos bioquímicos e morfológicos, com o intuito de ampliar o estudo das bases moleculares do HC e buscar um maior entendimento das causas genéticas das más formações da glândula. Optamos por utilizar a tecnologia de sequenciamento do exoma e realizamos uma análise exaustiva, tentando esgotar todas as possibilidades de análises para a seleção de variantes. Focamos a maior parte das nossas análises nas regiões codificadora e promotora dos genes com ênfase em variantes de troca de aminoácido ou pequenas deleções e inserções (indels). Também avaliamos a presença de regiões em homozigose, variações de número de cópias (CNVs) e sítios de ligação de microRNAs em sequências 3'UTRs. Obtivemos um sequenciamento homogêneo e de qualidade para todas as amostras. Cerca de 60% das bases tiveram leitura maior que 30X, foi possível alcançar uma boa cobertura para uma análise de precisão e confiabilidade do sequenciamento (Anexo 4).

Ao todo, identificamos 4.565 genes alterados em comum nos 35 pacientes, onde cerca da metade destes genes (52,4%) estavam alterados em todas as disgenesias e sinalizaram 38 termos envolvidos em eventos de diferenciação, proliferação, migração, adesão e morfogênese celular. Este primeiro dado sugere a existência de defeitos em genes envolvidos com o desenvolvimento, porém é de se notar que diversos outros termos não relacionados ao desenvolvimento também surgiram nesta análise. Este era um resultado esperado, devido ao grande número de genes inseridos na análise. Nosso objetivo com esta primeira análise foi ter uma visão geral dos genes alterados nos pacientes e seu envolvimento com a embriogênese. Além disso, também identificamos 6.830 variantes em comum nos 35 pacientes, e dentre elas 30 variantes raras e deletérias em 7 genes. Algumas destas alterações já eram descritas, porém sem dados na literatura sobre quadro clínico ou fenótipos relacionados à presença da variante. Esta foi uma etapa importante pois poderia trazer mais informações sobre estas alterações selecionadas. Verificamos que estes 7 genes não apresentavam interações de alta confiança, mostrando que não contribuem conjuntamente para uma função compartilhada. É de se notar que estes sete genes são muito polimórficos e nenhum deles já foi associado ao

desenvolvimento tireoidiano. Ainda assim, observamos 3 termos associados ao desenvolvimento que incluíram os genes *CDC27* e *IGSF3* (tabela). Um estudo recente mostrou que genes envolvidos na resposta imune, como o gene *IGSF3* (*Immunoglobulin Superfamily Member 3*) estão dentro dos mais polimórficos do genoma e que polimorfismos em regiões codificantes que afetam as sequências e/ou funções de proteínas são os menos frequentes no genoma humano (106). Estas informações junto com nossos resultados sugerem não haver uma relação das variantes destes genes com o fenótipo dos nossos pacientes. Por tanto, optamos por continuar nossas análises investigando alterações raras em genes diretamente relacionados com a tireoide. Contudo, os resultados deste estudo sobre genes alterados e variantes em comum ainda podem ser mais explorados com análises *in silico* e até ensaios funcionais para determinar seu envolvimento nas DTs.

Em nosso primeiro trabalho publicado com parte dos resultados deste estudo, onde relatamos o resultados do sequenciamento do exoma de dois tecidos ectópicos, descrevemos uma lista de genes candidatos para DT (107). Esta lista foi cuidadosamente criada coletando dados de trabalhos que já relacionaram mutações genéticas com o desenvolvimento ou função tireoidiana, dos mais antigos aos atuais (Anexo 5). Identificamos 750 variantes em 115 genes candidatos nos pacientes, e após aplicar os filtros de frequência e predição, chegamos ao número de 61 variantes raras e deletérias presentes em 42 genes candidatos (Tabela 4). A maioria destas variantes foi identificada em heterozigose e não observamos genes com várias alterações em diferentes pacientes ou genes com mais de uma alteração no mesmo paciente sugerindo uma heterozigose composta. Por tanto, podemos sugerir que a DT não teria um padrão de herança recessivo clássico. Algumas variantes identificadas são descritas na literatura em associação à fenótipos distintos de má formação crânio faciais (92-94), ao HC levando à disormonogênese por defeito na organificação do iodeto (95-97) e bócio (98, 99) e ainda em casos de câncer de tireoide (100), e não identificamos mutações ou outras variantes já relacionadas com a malformação tireoidiana.

Dentre os mecanismos conhecidos que podem levar às falhas de desenvolvimento tireoidiano está o modelo poligênico, onde falhas distintas em diferentes genes levam ao fenótipo de DT. Em modelo animal, ratos *Pax8 +/-* ou *Titf1*

+/- não apresentaram alteração no desenvolvimento da tireoide, enquanto ratos [Titf1, Pax8] +/- apresentaram glândulas tireoides menores ou hemiagenesia (8). O modelo poligênico é justificado pelos fatos de mutações monogênicas causais serem identificadas, até agora, em apenas uma pequena fração de pacientes com DT (5%), e que um padrão de herança mendeliano é raramente observado (108, 109). Além disso, já foi observada uma discordância de fenótipos entre irmãos portadores da mesma mutação, sugerindo o envolvimento conjunto de outros genes na determinação do defeito de desenvolvimento tireoidiano (42, 110). Como muitos dos pacientes apresentaram variantes raras e deletérias em mais de um gene candidato, verificamos a existência de interações entre estes genes alterados. Observamos interações de média confiança entre os genes candidatos alterados de 5 pacientes, mostrando que estes genes não participariam de uma mesma função.

Algumas das variantes em genes candidatos se destacaram, como a troca p.T574S no gene *TSHR* identificada no paciente 20 (hipoplasia). Não identificamos a alteração em 83 controles brasileiros normais, dentre estes, 33 são da mesma região do paciente, mostrando não haver alguma prevalência regional para a alteração. Esta variante não é descrita na literatura em associação à malformação tireoidiana ou outros fenótipos. O *TSHR* é um dos genes clássicos participantes da formação e função tireoidiana, sendo expresso durante o desenvolvimento da glândula (2). Nossas análises *in silico* mostram que a alteração é deletéria e afetaria a função da proteína. Também identificamos no paciente 20 a troca p.K101N no gene *FGF3*, que foi a única variante em homozigose observada na análise de genes candidatos. O *FGF3* faz parte da família FGF, proteínas que estão envolvidos em diversos processos biológicos, incluindo desenvolvimento embrionário, crescimento celular, morfogênese, reparo de tecidos, crescimento e invasão tumoral (*RefSeq*, Jul 2018). Os FGFs são expressos no desenvolvimento de todo o aparato faríngeo e sistema cardiovascular, sua ação seria via o fator de transcrição *TBX1* (111, 112). Também foi proposto que a redução de expressão de *Fgf3* em ratos aumentaria o fenótipo de haploinsuficiência de *Tbx1*, onde o *Fgf3* colaboraria para o desenvolvimento do endoderma faríngeo. Porém, foi observado que os defeitos mais severos não ocorreram na presença do *Fgf3* *-/-* (113). No paciente portador da alteração *FGF3*_p.K101N não identificamos variantes no gene *TBX1* ou variantes em outros genes FGF. As nossas análises *in*

silico mostraram que a variante apresenta potencial deletério e que não haveria interações entre o FGF3 e o TSHR.

Outra importante variante na análise de genes candidatos foi a troca p.S235X no gene *SFRP2*, identificada no paciente 35 (ectopia). O *SFRP2* é um dos moduladores da sinalização Wnt, participando do desenvolvimento de diversos tecidos (RefSeq, Julho 2008) e foi associado à ectopia estando diferentemente expresso em tecido tireoidiano (114). Por ser uma variante que causa um códon de parada prematuro, a função da proteína estaria automaticamente comprometida, o que também foi observado nas análises *in silico*. Sugerimos existir um significado clínico para esta variante que pode estar relacionada ao defeito de desenvolvimento do paciente. Assim, tanto a variante *SFRP2_p.S235X* no paciente 35 quanto as variantes *TSHR_p.T574S* e *FGF3_p.K101N* no paciente 20 poderiam comprometer a função das proteínas, esta predição deverá ser confirmada em análises funcionais futuras.

Diante destes primeiros resultados, voltamos nossas buscas para outras regiões que poderiam estar envolvidas na causa da doença dos nossos pacientes. Os FTs clássicos estão presentes desde os primeiros estágios da embriogênese tireoidiana (7), contudo, não identificamos estes FTs afetados por variantes raras e deletérias. A expressão de genes é controlada através da ligação dos FTs às sequências *motifs* presentes nos promotores (115). Não identificamos variantes raras em *motifs* de ligação de FOXE1 e PAX8, indicando que a ação destes FTs não estaria comprometida. Contudo, identificamos variantes raras e deletérias na região codificadora de genes regulados por FOXE1 e PAX8, com a hipótese de que FOXE1 e PAX8 estariam se ligando corretamente aos respectivos *motifs*, porém a codificação dos genes estaria comprometida. Algumas variantes identificadas são descritas na literatura em associação à fenótipos distintos de síndrome de Walker-Warburg e distrofia muscular (101-103), e de HC levando à disormonogênese por defeito na organificação do iodeto (95-97) e bócio (98, 99). As análises de enriquecimento e interação destes genes alterados regulados por FOXE1 e PAX8 mostraram não haver um envolvimento dos genes alterados com o desenvolvimento tireoidiano. Como anteriormente nas análises de genes candidatos, não observamos mutações ou outras variantes já relacionadas com a malformação tireoidiana nesta análise. É

importante destacar que não descartamos a importância destes achados na investigação sobre a doença dos pacientes. Análises futuras poderão esclarecer o significado destas alterações e trazer mais informações sobre os mecanismos do desenvolvimento tireoidiano.

Apesar do papel de FOXE1 no desenvolvimento e função da tireoide já ser bem conhecido, poucas mutações causadoras de doenças já foram descritas no gene (116-122). Nossa investigação do exoma das 35 DTs não revelou variantes raras e deletérias nas regiões codificadora e regulatórias do gene *FOXE1*. Entretanto, identificamos 25 SNVs e 1 indel frequentes em diferentes populações, dentre elas as variantes rs1867277 e rs71369530. A variante rs1867277 está localizada na região promotora de FOXE1. Foi proposto anteriormente um modelo de regulação de *FOXE1* dependente do genótipo de rs1867277, onde a presença de um alelo A aumentaria a suscetibilidade ao CT através do recrutamento dos FTs USF1/USF2 (123). Essa regulação alelo específica aumentaria a atividade transcricional do promotor de *FOXE1*. Da mesma forma, esse polimorfismo funcional foi associado como um fator genético causador de fenda labiopalatina em um estudo de casos familiares (124). Estas evidências nos levaram a estender a investigação do rs1867277 para mais pacientes com DT em nossa coorte e em indivíduos normais brasileiros sem doença tireoidiana. Nossas análises mostraram uma forte associação do alelo A de rs1867277 com a presença de DT em nossos pacientes, com uma maior predisposição tanto na presença do genótipo dominante quanto do recessivo. O polimorfismo rs1867277 também foi descrito em desequilíbrio de ligação com a variante rs907577 (123), localizado *upstream* à região promotora de *FOXE1* e incluído dentro de um bloco de risco para hipotireoidismo primário (104). Esta variante não foi identificada no sequenciamento do exoma, por tanto foi necessário realizar sua genotipagem. Contudo, a presença do alelo menor C de rs907577 não demonstrou correlação com o defeito de desenvolvimento da tireoide dos nossos pacientes. Outros dois outros estudos investigaram a presença da variante rs907577 e não encontraram associação com fissuras orofaciais na população chinesa (124, 125).

A variante rs71369530 afeta o comprimento da cadeia polialanina de FOXE1 (a maioria das contagens de alelos relatados são 16^{Ala} e 14^{Ala}). A contagem de 14^{Ala} na cadeia polialanina de FOXE1 foi associada a um risco aumentado na ocorrência

de DT, com uma associação mais forte com o grupo de ectopias. Por outro lado, a presença de 16^{Ala} teria um efeito protetor contra a doença, sugerindo que rs71369530 pode ser uma variante chave durante a embriogênese tireoidiana (126, 127). Da mesma forma, foi demonstrado a associação da contagem de alaninas com a ocorrência de CT, entretanto, neste transtorno a contagem de 16^{Ala} foi considerada um fator de risco (128). Assim, o padrão de 16^{Ala} afetaria inversamente o CT e os defeitos de desenvolvimento da tireoide. A contagem de 16^{Ala} da rs71369530 também foi descrita em desequilíbrio de ligação com a variante rs1867277 em pacientes com CT papilífero, sem ocasionar diferenças na expressão da proteína nem na ligação ao DNA (129). Esta associação do tamanho da cadeia de polialanina e o CT papilífero não seria autônoma, dependendo da presença de outros fatores de predisposição, como rs1867277 (130). Esses achados revelaram um duplo papel de FOXE1 desde o início da formação do primórdio tireoidiano até a ocorrência de doenças malignas da tireoide. Estas informações nos levaram a também incluir o SNP rs71369530 em nossas análises. Observamos a associação do alelo menor de rs71369530 (14^{Ala}) com a DT na comparação com *ABraOM*, e o mesmo resultado observado na comparação do padrão de herança recessivo. No entanto, a comparação com o nosso grupo CN não demonstrou os mesmos resultados, sem associação e não aumentando as chances de desenvolvimento da doença em ambas as análises (padrões de alelos e genótipos). É de se notar que o grupo CN e a base de dados *ABraOM* são compostos por dados de indivíduos brasileiros. O *ABRAOM* inclui dados de maiores de 60 anos, da cidade de São Paulos, sem doenças de base, mas não possui informações sobre a função tireoidiana destes indivíduos (76). Nosso grupo CN é mais heterogêneo em idade e inclui indivíduos das cidades de São Paulo (como a coorte *ABraOM*) e Marília (pais de pacientes) no estado de São Paulo e da cidade de Salvador (como a maioria dos nossos casos de DT) na Bahia. A população brasileira tem origem diversa, o que afeta a ancestralidade genômica e pode causar prevalências regionais envolvendo diferenças genotípicas. A distribuição da população no vasto território brasileiro, fatores ambientais e até mesmo o desenvolvimento econômico já foram apontados como os principais fatores que seriam responsáveis pela heterogênea genética brasileira (131, 132). Especulamos a existência de diferenças regionais que levam a alguma prevalência genética envolvendo nosso grupo CN e a população do *ABraOM*,

assim como não descartamos o fato de que as diferenças no número de indivíduos em cada grupo controle podem ter afetado os resultados.

Após observar a associação individual das variantes rs1867277 e rs71369530 com DT, também verificamos que a combinação dos genótipos alterados das duas variantes demonstrou uma associação ainda maior (Tabela 10). Nikitski et al. descreveram a ação independente do alelo A de rs1867277 no CT papilar, enquanto o efeito de 16^A de rs71369530 dependeria da presença de outros fatores (130). Os autores propuseram que o papel deletério de rs71369530 seria causado por sua presença em um forte desequilíbrio de ligação com outras variantes de risco, como a rs1867277. Considerando este mesmo modelo, já que em nossos casos observamos a associação independente de rs1867277 e de rs71369530 e uma associação mais forte na avaliação conjunta dessas duas variantes com a DT, investigamos o efeito de outras variantes dentro do mesmo bloco cromossômico. Utilizando os dados do sequenciamento do exoma das 35 DTs, identificamos outras 4 variantes dentro desta região (rs7849497, rs1867278, rs1867279 e rs1867280) que exibiram associação individual com as DT na comparação com *ABraOM*. A alteração rs7849497 não possui informações da literatura, apenas informações de frequência nas bases de dados. Portanto, não foi possível reunir informações genéticas e fenotípicas sobre as causas da presença do alelo menor desta variante em estudos já publicados. As alterações rs1867277, rs1867278, rs1867279 e rs1867280 estão dentro de um bloco de desequilíbrio de ligação no cromossomo 9q22 já relacionado com malformação de lábio leporino com ou sem fenda palatina (CL/P) (133, 134) e com CT (123, 129, 135, 136). A variante rs1867278 foi avaliada em um estudo familiar e apresentou a maior associação com CL/P (133). As variantes rs1867279 e rs1867280 não foram identificadas em CL/P com casos de HC ou CT, mas análises *in silico* mostraram que ambas as alterações podem modular a expressão de genes relevantes envolvidos na morfogênese tireoidiana por meio da modificação dos sítios de ligação do fator de transcrição. No entanto, os autores propõem um modelo de resposta disruptiva em que etnia, fatores ambientais e os dois SNPs atuam juntos para produzir o fenótipo (137). Apesar de observar essa associação significativa das 6 alterações no exoma de 35 casos de DT, estes resultados não se repetiram quando se comparou com população do 1000G. Especulamos que este resultado pode ser uma consequência da grande diferença entre o número de indivíduos, 35 DTs contra 2504 indivíduos no

banco de dados 1000G. Importante lembrar que nossas análises anteriores que associaram as variantes rs1867277 e rs71369530 utilizando o banco de dados 1000G incluíram toda a casuística de 71 casos de DT. Esses resultados sugerem que não apenas as variantes rs1867277 e rs71369530 podem estar envolvidas na predisposição para malformações da tireoide, mas também variantes presentes dentro deste bloco no cromossomo 9q22, como em estudos anteriores que mostraram que variantes na região do promotor de *FOXE1* estavam associadas com a ocorrência de HC e com CT não medular familiar (104, 138). Um estudo de desequilíbrio de ligação com maior quantidade de amostras pode fornecer mais informações sobre a associação de todas essas variantes no gene *FOXE1* com os defeitos na embriogênese da tireoide.

Para completar as análises de *FOXE1*, avaliamos as variantes na região 3'UTR e verificamos que nenhum dos sítios conservados de ligação de miRNAs na região 3'UTR de *FOXE1* foi alterado ou comprometido. Os miRNAs desempenham um papel importante durante a embriogênese e estudos têm mostrado que a desregulação de miRNAs específicos pode levar a erros no desenvolvimento da glândula tireoide, com ruptura do tecido tireoidiano e poucas estruturas foliculares (139, 140). Para que a regulação mediada por miRNAs ocorra, é essencial o pareamento da sequência *seed* dos miRNAs com as regiões complementares presentes, na maioria dos casos, nas regiões 3'UTR do RNAm (141). Regiões parcialmente complementares são muito comuns em animais, mas impedem a regulação por miRNAs (142). Com estes dados, as variantes presentes na região 3'UTR de nossos pacientes não estariam associadas à causa da DT.

Após a introdução das tecnologias de sequenciamento de nova geração, diversos trabalhos vêm mostrando que a malformação tireoidiana também pode estar associada à presença de variantes em genes relacionados à síntese hormonal tireoidiana. Já foram descritas mutações no gene *TPO* como causa de hipoplasia tireoidiana, sendo que a proteína é responsável pela etapa de ligação do iodo à tireoglobulina na produção dos hormônios tireoidianos (52). Wang et al. mostrou os genes *DUOX* como patogênicos, sendo *DUOX2* o gene variante mais comum em disgenesias e *DUOX1* o gene variante mais comum em casos de glândula *in situ* (105). O estudo brasileiro conduzido por Kizys et al, descreveu uma extensa análise do gene

DUOX2 em uma coorte de pacientes com DT (53). O trabalho utilizou da tecnologia de WES e buscou avaliar a causa da DT nos pacientes, finalizando com a descrição de diversas variantes no *DUOX2* com a função deletéria em ectopias tireoidianas. Em nossos pacientes, também identificamos variantes deletérias no gene *DUOX2* e as análises *in silico* mostraram que todas seriam prejudiciais, implicando em alterações nas interações moleculares ou modificando a estrutura tridimensional da proteína. Isso nos mostra que estas variantes podem ser relevantes e um estudo funcional está em andamento para fornecer mais informações sobre a ação dessas alterações e uma possível relação na causa da doença dos pacientes.

Para completar as análises procuramos por regiões cromossômicas alteradas. A dissomia uniparental parcial é a herança de duas cópias de uma região cromossômica de um mesmo genitor, que gera regiões em homozigose e pode levar à expressão de variantes deletérias (143). Utilizando os dados do sequenciamento do exoma identificamos duas regiões em homozigose em dois pacientes, porém sem variantes raras dentro das regiões. Dessa forma, a presença de homozigosidade não foi associada às DTs neste estudo.

Também analisamos a presença de regiões cromossômicas duplicadas ou deletadas. As CNVs têm surgido como um componente genético importante de doenças, como câncer, autismo, doenças do desenvolvimento, entre outras (54, 55). Nos últimos anos, foram desenvolvidas novas ferramentas que permitem identificar CNVs através dos dados do sequenciamento de exoma. Utilizando esta linha de análises, identificamos 10 CNVs presentes em 12 dos nossos pacientes, sendo que 2 destas CNVs (ambas duplicações) envolviam a sequência codificante de genes (*LOC283177* e *AHNAK2*), não relacionados com a embriogênese da tireoide. Este resultado está de acordo com a literatura, que indicam que as CNVs não são os eventos genéticos principais das causas da DT (54, 114).

Por fim, realizamos uma última análise avaliando todas as informações sobre variantes que identificamos nas etapas anteriores. O objetivo desta etapa foi correlacionar as variantes em diferentes genes com informações contidas na literatura e com o fenótipo e dados bioquímicos dos pacientes. Ao reunir todos os achados de

cada exoma, verificamos que 97,1% dos pacientes possuíam variantes em genes diretamente relacionados com o desenvolvimento e/ou função tireoidiana.

Não observamos qualquer prevalência de fenótipos ou genes alterados, assim como não havia correlação com os dados bioquímicos, reforçando a inexistência de um padrão genético para as causas da DT, o que já foi observado em morfologias distintas de desenvolvimento tireoidiano, como agenesia e ectopia, sendo representadas pela mesma causa genética (10, 144). Num modelo poligênico, podemos considerar que a combinação de variantes em diferentes genes pode estar relacionada com o desenvolvimento de DT em nossos casos. Verificamos esta hipótese com as análises de interação, onde observamos interações de alta confiança. Dentre estas interações, algumas foram observadas entre genes que já foram descritos na morfogênese tireoidiana, como no paciente 3 (agenesia) que possui variantes nos genes *PAX3* e *ISL1* com uma pontuação com tendência a alta confiança. A expressão do gene *ISL1* foi descrita como presente desde a etapa de diferenciação e brotamento, até a migração e as fases finais do desenvolvimento tireoidiano (145). O gene *PAX3* é potencialmente envolvido com defeitos de desenvolvimento que levam à hipoplasia tireoidiana (2), e apesar do paciente possuir ausência do tecido, não descartamos o envolvimento conjunto das variantes nestes genes na presença das DT. Já no paciente 33 (ectopia) verificamos uma interação cíclica entre os genes *JAG1*, *RUN2* e *BMP4* de média a alta confiança. Os genes *RUNX2* e *BMP4* estão envolvidos com o desenvolvimento de diversos outros tecidos (Ensembl). O *BMP4* é um ligante do receptor BMP ativando a via BMP, que recentemente se mostrou extremamente necessária durante os primeiros estágios do desenvolvimento do endoderma cardíaco e tireoidiano (146). Já o gene *Jag1* seria necessário durante a proliferação e manutenção do estado diferenciado das células tireoidianas em modelo de zebrafish, atuando como um fator de predisposição para vários defeitos de desenvolvimento tireoidiano (147). Os resultados de interação demonstram que os três genes participariam de uma mesma função compartilhada, e as variantes poderiam estar envolvidas no defeito do paciente.

Também observamos interações de alta confiança entre o gene *FOXE1* e outros genes. No paciente 20 (hipoplasia) observamos a interação entre os genes *FOXE1* e *TSHR*, com pontuação de alta confiança. O *TSHR* é expresso durante o estágio de expansão das células diferenciadas da tireoide (2). As variantes em ambos

os genes se mostraram deletérias e associadas com a DT, o que poderiam contribuir para o fenótipo do paciente. Também consideramos que as interações entre *FOXE1* e *TPO* no paciente 1 (agenesia) e *TG* nos pacientes 2 (agenesia) e 34 (ectopia) poderiam estar associadas à malformação tireoidiana. O *TPO* é um dos genes mais alterados em HC severo e DT (148), sendo que casos de hipoplasia tireoidiana já foram associados a variantes neste gene (52). Da mesma maneira, alterações no gene *TG* também foram identificadas em casos de TD, porém em menor prevalência (149). A interação de *TPO* e *TG* com o fator de transcrição *FOXE1* se dá também porque *FOXE1* está presente não apenas no desenvolvimento mas também na manutenção do tecido tireoidiano, indispensável também para a função da glândula (2).

Finalmente, verificamos uma interação de alta confiança entre os genes *SPARC* e *COL1A1*, alterados no paciente 31 (ectopia). O *SPARC* não está envolvido diretamente com o desenvolvimento tireoidiano, foi descrito como regulador da proliferação de células endoteliais (150) e participa do desenvolvimento de outros tecidos, como coração, ouvido interno e pulmão (Ensembl). O gene *COL1A1* já foi associado ao câncer e teria influência na invasão do câncer de tireoide, atuando na progressão do CT papilífero (151). Hipotetizamos que o mesmo efeito de invasão no CT possa estar envolvido no movimento defeituoso de migração participando da formação da ectopia tireoidiana no paciente 31. A interação de maior confiança foi observada no paciente 15 (hipoplasia) entre os genes *DUOX2* e *DUOXA1*. As proteínas *DUOX* geram o H_2O_2 necessário para a organificação do iodeto na hormonogênese tireoidiana (30), e os genes da família *DUOX* se localizam no cromossomo 15 de maneira contígua e cada uma das isoformas *DUOX/DUOXA* compartilham do mesmo promotor bidirecional (152). Logo, é esperado uma alta confiança de interação entre os genes. É importante considerar que apenas os complexos *DUOX1/DUOXA1* e *DUOX2/DUOXA2* são funcionais, sendo o complexo *DUOX2/DUOXA2* o mais expresso na tireoide e com maior eficiência na produção de H_2O_2 (153). Apesar de sua conhecida ação na hormonogênese, diversos estudos mostram os genes *DUOX2* e *DUOXA2* alterados com alta frequência em pacientes com DT (53, 105, 148, 149), principalmente em casos com glândula in situ (105), como no paciente 15.

Em resumo, nossas análises associaram variantes no gene *FOXE1* ao defeito de DT e identificaram genes e variantes potencialmente importantes para o desenvolvimento da tireoide. Não descartamos a necessidade de ensaios funcionais e até mesmo o uso de modelos animais que podem fornecer dados mais conclusivos sobre estes resultados. Alguns ensaios já estão em andamento junto ao nosso grupo de pesquisa, e ainda é possível reanalisar os dados dos exomas dos pacientes com novas ferramentas que surjam e assim trazer novos dados sobre a genética da embriogênese tireoidiana. Os resultados de interação observados reforçam a necessidade de uma pesquisa mais ampla de pacientes com DT, onde o foco durante muitos anos foi apenas a pesquisa por alterações monogênicas, agora considerando uma possível doença poligênica. Ainda deve ser considerado que outros mecanismos podem estar envolvidos, como mutações somáticas precoces, eventos pós zigóticos e modificações epigenéticas (2, 43, 154).

6 CONCLUSÕES

Este estudo permitiu um amplo rastreamento de variantes em casos de DT e identificar variantes com potencial relevância clínica nestes pacientes. Os resultados sugerem que genes até então não associados às DTs podem estar envolvidos na doença. Os polimorfismos no gene *FOXE1* foram fatores de risco para o desenvolvimento de malformações da tireoide. Os dados também reforçam a inexistência de um padrão genético único para as causas da DT e, além disso, que os modelos poligênico ou de múltiplos fatores podem estar relacionados com as falhas de desenvolvimento da tireoide, os quais devem ser mais explorados. Não descartamos a necessidade de ensaios funcionais e até mesmo o uso de modelos animais que podem fornecer dados mais conclusivos a associação destas variantes com a DT.

7 REFERÊNCIAS

1. Maciel RMB. Tireoide: fisiologia e avaliação diagnóstica. *Endocrinologia*. 1. São Paulo: Ed. Atheneu; 2007. p. 299-329.
2. De Felice M, Di Lauro R. Thyroid development and its disorders: genetics and molecular mechanisms. *Endocr Rev*. 2004;25(5):722-46.
3. O'Rahilly R. The timing and sequence of events in the development of the human endocrine system during the embryonic period proper. *Anat Embryol (Berl)*. 1983;166(3):439-51.
4. Fagman H, Nilsson M. Morphogenesis of the thyroid gland. *Mol Cell Endocrinol*. 2010;323(1):35-54.
5. Postiglione MP, Parlato R, Rodriguez-Mallon A, Rosica A, Mithbaokar P, Maresca M, et al. Role of the thyroid-stimulating hormone receptor signaling in development and differentiation of the thyroid gland. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99(24):15462-7.
6. Szinnai G. Genetics of normal and abnormal thyroid development in humans. *Best Pract Res Clin Endocrinol Metab*. 2014;28(2):133-50.
7. Fernández LP, López-Márquez A, Santisteban P. Thyroid transcription factors in development, differentiation and disease. *Nat Rev Endocrinol*. 2015;11(1):29-42.
8. Amendola E, De Luca P, Macchia PE, Terracciano D, Rosica A, Chiappetta G, et al. A mouse model demonstrates a multigenic origin of congenital hypothyroidism. *Endocrinology*. 2005;146(12):5038-47.
9. Parlato R, Rosica A, Rodriguez-Mallon A, Affuso A, Postiglione MP, Arra C, et al. An integrated regulatory network controlling survival and migration in thyroid organogenesis. *Dev Biol*. 2004;276(2):464-75.
10. De Felice M, Ovitt C, Biffali E, Rodriguez-Mallon A, Arra C, Anastassiadis K, et al. A mouse model for hereditary thyroid dysgenesis and cleft palate. *Nature genetics*. 1998;19(4):395-8.
11. Antonica F, Kasprzyk DF, Opitz R, Iacovino M, Liao XH, Dumitrescu AM, et al. Generation of functional thyroid from embryonic stem cells. *Nature*. 2012;491(7422):66-71.
12. Wendl T, Adzic D, Schoenebeck JJ, Scholpp S, Brand M, Yelon D, et al. Early developmental specification of the thyroid gland depends on hox-expressing surrounding tissue and on FGF signals. *Development*. 2007;134(15):2871-9.
13. Lania G, Zhang Z, Huynh T, Caprio C, Moon AM, Vitelli F, et al. Early thyroid development requires a Tbx1-Fgf8 pathway. *Developmental biology*. 2009;328(1):109-17.

14. Simpser T, Rapaport R. Update on some aspects of neonatal thyroid disease. *J Clin Res Pediatr Endocrinol*. 2010;2(3):95-9.
15. Setian NS. Hypothyroidism in children: diagnosis and treatment. *J Pediatr (Rio J)*. 2007;83(5 Suppl):S209-16.
16. Porterfield SP, Hendrich CE. The role of thyroid hormones in prenatal and neonatal neurological development--current perspectives. *Endocr Rev*. 1993;14(1):94-106.
17. Loeber JG. Neonatal screening in Europe; the situation in 2004. *J Inherit Metab Dis*. 2007;30(4):430-8.
18. Deladoëy J, Pfarr N, Vuissoz JM, Parma J, Vassart G, Biesterfeld S, et al. Pseudodominant inheritance of goitrous congenital hypothyroidism caused by TPO mutations: molecular and in silico studies. *J Clin Endocrinol Metab*. 2008;93(2):627-33.
19. 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(5):739-45.
20. Maciel LM, Kimura ET, Nogueira CR, Mazeto GM, Magalhaes PK, Nascimento ML, et al. Congenital hypothyroidism: recommendations of the Thyroid Department of the Brazilian Society of Endocrinology and Metabolism. *Arq Bras Endocrinol Metabol*. 2013;57(3):184-92.
21. Nunes AK, Wachholz RG, Rover MR, Souza LC. [Prevalence of disorders detected by newborn screening in Santa Catarina]. *Arq Bras Endocrinol Metabol*. 2013;57(5):360-7.
22. Mendes LC, Santos TT, Bringel FeA. [Evolution of the neonatal screening program in the state of Tocantins]. *Arq Bras Endocrinol Metabol*. 2013;57(2):112-9.
23. Matos DM, Ramalho RJ, Carvalho BM, Almeida MA, Passos LF, Vasconcelos TT, et al. Evolution to permanent or transient conditions in children with positive neonatal TSH screening tests in Sergipe, Brazil. *Arch Endocrinol Metab*. 2016;60(5):450-6.
24. Palhares HM, Silva LC, Sato LM, Lara BH, Miranzi SeS, Silva AeP, et al. Incidence of congenital hypothyroidism in the city of Uberaba/Minas Gerais and etiological evaluation of the affected subjects. *Arq Bras Endocrinol Metabol*. 2012;56(5):305-12.
25. de Vijlder JJ, Ris-Stalpers C, Vulsma T. Inborn errors of thyroid hormone biosynthesis. *Exp Clin Endocrinol Diabetes*. 1997;105 Suppl 4:32-7.
26. Beltrão CB, Juliano AG, Chammas MC, Watanabe T, Sapienza MT, Marui S. Etiology of congenital hypothyroidism using thyroglobulin and ultrasound combination. *Endocr J*. 2010;57(7):587-93.

27. Ramos JC, de Lacerda Filho L, Demartini AeA, da Silveira RB, Pereira RM, Sandrini Neto R, et al. Clinical and laboratory features of children and adolescents with congenital hypothyroidism due to dysmorphogenesis in southern Brazil. *Arq Bras Endocrinol Metabol.* 2012;56(3):201-8.
28. Spitzweg C, Morris JC. Genetics and phenomics of hypothyroidism and goiter due to NIS mutations. *Mol Cell Endocrinol.* 2010;322(1-2):56-63.
29. Bizhanova A, Kopp P. Genetics and phenomics of Pendred syndrome. *Mol Cell Endocrinol.* 2010;322(1-2):83-90.
30. Grasberger H. Defects of thyroidal hydrogen peroxide generation in congenital hypothyroidism. *Mol Cell Endocrinol.* 2010;322(1-2):99-106.
31. Nascimento AC, Guedes DR, Santos CS, Knobel M, Rubio IG, Medeiros-Neto G. Thyroperoxidase gene mutations in congenital goitrous hypothyroidism with total and partial iodide organification defect. *Thyroid.* 2003;13(12):1145-51.
32. Pardo V, Vono-Toniolo J, Rubio IG, Knobel M, Possato RF, Targovnik HM, et al. 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.* 2009;94(8):2938-44.
33. Moreno JC, Visser TJ. Genetics and phenomics of hypothyroidism and goiter due to iodotyrosine deiodinase (DEHAL1) gene mutations. *Mol Cell Endocrinol.* 2010;322(1-2):91-8.
34. Setian N. Tireóide. In: Setian, editor. *Endocrinologia Pediátrica: aspectos físicos e metabólicos do recém-nascido ao adolescente* São Paulo: Sarvier; 2002.
35. De Felice M, Di Lauro R. Thyroid development and its disorders: genetics and molecular mechanisms. *Endocr Rev.* 2004;25(5):722-46.
36. Ramos HE, Nesi-França S, Maciel RM. [New aspects of genetics and molecular mechanisms on thyroid morphogenesis for the understanding of thyroid dysgenesis]. *Arq Bras Endocrinol Metabol.* 2008;52(9):1403-15.
37. De Felice M, Di Lauro R. Minireview: Intrinsic and extrinsic factors in thyroid gland development: an update. *Endocrinology.* 2011;152(8):2948-56.
38. Knobel M, Medeiros-Neto G. An outline of inherited disorders of the thyroid hormone generating system. *Thyroid.* 2003;13(8):771-801.
39. Arancibia P, Veliz J, Barria M, Pineda G. Lingual thyroid: report of three cases. *Thyroid.* 1998;8(11):1055-7.
40. Kreisner E, Camargo-Neto E, Maia CR, Gross JL. Accuracy of ultrasonography to establish the diagnosis and aetiology of permanent primary congenital hypothyroidism. *Clinical endocrinology.* 2003;59(3):361-5.
41. Castanet M, Polak M, Leger J. Familial forms of thyroid dysgenesis. *Endocr Dev.* 2007;10:15-28.

42. Castanet M, Marinovic D, Polak M, Leger J. Epidemiology of thyroid dysgenesis: the familial component. *Horm Res Paediatr*. 2010;73(4):231-7.
43. Perry R, Heinrichs C, Bourdoux P, Khoury K, Szots F, Dussault JH, et al. Discordance of monozygotic twins for thyroid dysgenesis: implications for screening and for molecular pathophysiology. *J Clin Endocrinol Metab*. 2002;87(9):4072-7.
44. Olivieri A, Medda E, De Angelis S, Valensise H, De Felice M, Fazzini C, et al. High risk of congenital hypothyroidism in multiple pregnancies. *J Clin Endocrinol Metab*. 2007;92(8):3141-7.
45. 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. *The Journal of clinical endocrinology and metabolism*. 2010;95(4):1981-5.
46. Montanelli L, Tonacchera M. Genetics and phenomics of hypothyroidism and thyroid dys- and agenesis due to PAX8 and TTF1 mutations. *Mol Cell Endocrinol*. 2010;322(1-2):64-71.
47. Mio C, Grani G, Durante C, Damante G. Molecular defects in thyroid dysgenesis. *Clin Genet*. 2020;97(1):222-31.
48. Fagman H, Liao J, Westerlund J, Andersson L, Morrow BE, Nilsson M. The 22q11 deletion syndrome candidate gene *Tbx1* determines thyroid size and positioning. *Hum Mol Genet*. 2007;16(3):276-85.
49. Muzza M, Persani L, de Filippis T, Gastaldi R, Vigone MC, Sala D, et al. Absence of sonic hedgehog (*Shh*) germline mutations in patients with thyroid dysgenesis. *Clin Endocrinol (Oxf)*. 2008;69(5):828-9.
50. Ferrara AM, Rossi G, Zampella E, Di Candia S, Pagliara V, Nettore IC, et al. Screening for mutations in the *ISL1* gene in patients with thyroid dysgenesis. *J Endocrinol Invest*. 2011;34(7):e149-52.
51. Kuhnen P, Turan S, Frohler S, Guran T, Abali S, Biebermann H, et al. Identification of *PENDRIN* (*SLC26A4*) mutations in patients with congenital hypothyroidism and "apparent" thyroid dysgenesis. *The Journal of clinical endocrinology and metabolism*. 2014;99(1):E169-76.
52. Stoupa A, Chaabane R, Guériouz M, Raynaud-Ravni C, Nitschke P, Bole-Feysot C, et al. Thyroid Hypoplasia in Congenital Hypothyroidism Associated with Thyroid Peroxidase Mutations. *Thyroid*. 2018;28(7):941-4.
53. Kizys MML, Louzada RA, Mitne-Neto M, Jara JR, Furuzawa GK, de Carvalho DP, et al. *DUOX2* Mutations Are Associated With Congenital Hypothyroidism With Ectopic Thyroid Gland. *J Clin Endocrinol Metab*. 2017;102(11):4060-71.
54. Thorwarth A, Mueller I, Biebermann H, Ropers HH, Grueters A, Krude H, et al. Screening chromosomal aberrations by array comparative genomic hybridization in 80

patients with congenital hypothyroidism and thyroid dysgenesis. *J Clin Endocrinol Metab.* 2010;95(7):3446-52.

55. Kim HJ, Jo HS, Yoo EG, Chung IH, Kim SW, Lee KH, et al. 22q11.2 Microduplication with thyroid hemiagenesis. *Hormone research in paediatrics.* 2013;79(4):243-9.

56. Schell LM, Gallo MV, Ravenscroft J. Environmental influences on human growth and development: historical review and case study of contemporary influences. *Ann Hum Biol.* 2009;36(5):459-77.

57. Deladoëy J, Bélanger N, Van Vliet G. Random variability in congenital hypothyroidism from thyroid dysgenesis over 16 years in Québec. *J Clin Endocrinol Metab.* 2007;92(8):3158-61.

58. McNally RJQ, Jones JH, Shaikh MG, Donaldson MDC, Blakey K, Cheetham TD. Congenital Hypothyroidism: Space-Time Clustering of Thyroid Dysgenesis Indicates a Role for Environmental Factors in Disease Etiology. *Thyroid.* 2020.

59. Hashemipour M, Kelishadi R, Amin MM, Poursafa P, Rashidi M, Mehrnejat N, et al. The association between familial and environmental factors and prevalence of congenital hypothyroidism in center of Iran. *Environ Sci Pollut Res Int.* 2021;28(7):8434-41.

60. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America.* 1977;74(12):5463-7.

61. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet.* 2008;9:387-402.

62. Guan YF, Li GR, Wang RJ, Yi YT, Yang L, Jiang D, et al. Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer. *Chin J Cancer.* 2012;31(10):463-70.

63. Yu Y, Wu BL, Wu J, Shen Y. Exome and whole-genome sequencing as clinical tests: a transformative practice in molecular diagnostics. *Clin Chem.* 2012;58(11):1507-9.

64. Singleton AB. Exome sequencing: making hay while the sun shines. *Lancet Neurol.* 2011;10(10):942-6.

65. Singleton AB. Exome sequencing: a transformative technology. *Lancet Neurol.* 2011;10(10):942-6.

66. Marian AJ. Challenges in medical applications of whole exome/genome sequencing discoveries. *Trends Cardiovasc Med.* 2012;22(8):219-23.

67. Goh G, Choi M. Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. *Genomics Inform.* 2012;10(4):214-9.

68. Carré A, Stoupa A, Kariyawasam D, Gueriouz M, Ramond C, Monus T, et al. Mutations in BOREALIN cause thyroid dysgenesis. *Hum Mol Genet.* 2017;26(3):599-610.
69. Choukair D, Eberle B, Vick P, Hermanns P, Weiss B, Paramasivam N, et al. Identification of Transient Receptor Potential Channel 4-Associated Protein as a Novel Candidate Gene Causing Congenital Primary Hypothyroidism. *Horm Res Paediatr.* 2020;93(1):16-29.
70. Szczepanek-Parulska E, Budny B, Borowczyk M, Zawadzka K, Sztromwasser P, Ruchała M. Compound heterozygous GLI3 variants in siblings with thyroid hemiagenesis. *Endocrine.* 2021;71(2):514-9.
71. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16(3):1215.
72. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-60.
73. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-303.
74. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
75. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.
76. Naslavsky MS, Yamamoto GL, de Almeida TF, Ezquina SAM, Sunaga DY, Pho N, et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum Mutat.* 2017;38(7):751-63.
77. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018;46(D1):D1062-D7.
78. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29(1):308-11.
79. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res.* 2017;45(D1):D840-D5.
80. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* 2014;42(Database issue):D986-92.
81. Consortium U. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49(D1):D480-D9.

82. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016;44(W1):W90-7.
83. González-Pérez A, López-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet.* 2011;88(4):440-9.
84. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat.* 2013;34(1):57-65.
85. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics.* 2010;11:548.
86. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* 2011;39(17):e118.
87. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014;11(4):361-2.
88. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248-9.
89. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43(Database issue):D447-52.
90. Tsunoda T, Takagi T. Estimating transcription factor bindability on DNA. *Bioinformatics.* 1999;15(7-8):622-30.
91. Serra RW, Fang M, Park SM, Hutchinson L, Green MR. A KRAS-directed transcriptional silencing pathway that mediates the CpG island methylator phenotype. *Elife.* 2014;3:e02313.
92. Bertolacini CD, Ribeiro-Bicudo LA, Petrin A, Richieri-Costa A, Murray JC. Clinical findings in patients with GLI2 mutations--phenotypic variability. *Clin Genet.* 2012;81(1):70-5.
93. Suzuki S, Marazita ML, Cooper ME, Miwa N, Hing A, Jugessur A, et al. Mutations in BMP4 are associated with subepithelial, microform, and overt cleft lip. *Am J Hum Genet.* 2009;84(3):406-11.
94. Gregory LC, Gaston-Massuet C, Andoniadou CL, Carreno G, Webb EA, Kelberman D, et al. The role of the sonic hedgehog signalling pathway in patients with midline defects and congenital hypopituitarism. *Clin Endocrinol (Oxf).* 2015;82(5):728-38.

95. Moreno JC, Bikker H, Kempers MJ, van Trotsenburg AS, Baas F, de Vijlder JJ, et al. Inactivating mutations in the gene for thyroid oxidase 2 (THOX2) and congenital hypothyroidism. *N Engl J Med*. 2002;347(2):95-102.
96. Varela V, Rivolta CM, Esperante SA, Gruñeiro-Papendieck L, Chiesa A, Targovnik HM. Three mutations (p.Q36H, p.G418fsX482, and g.IVS19-2A>C) in the dual oxidase 2 gene responsible for congenital goiter and iodide organification defect. *Clin Chem*. 2006;52(2):182-91.
97. De Marco G, Agretti P, Montanelli L, Di Cosmo C, Bagattini B, De Servi M, et al. Identification and Functional Analysis of Novel Dual Oxidase 2 (DUOX2) Mutations in Children with Congenital or Subclinical Hypothyroidism. *J Clin Endocrinol Metab*. 2011;96(8):E1335-E9.
98. Corral J, Martín C, Pérez R, Sánchez I, Mories MT, San Millan JL, et al. Thyroglobulin gene point mutation associated with non-endemic simple goitre. *Lancet*. 1993;341(8843):462-4.
99. Vono-Toniolo J, Rivolta CM, Targovnik HM, Medeiros-Neto G, Kopp P. Naturally occurring mutations in the thyroglobulin gene. *Thyroid*. 2005;15(9):1021-33.
100. Schechter RB, Nagilla M, Joseph L, Reddy P, Khattri A, Watson S, et al. Genetic profiling of advanced radioactive iodine-resistant differentiated thyroid cancer and correlation with axitinib efficacy. *Cancer Lett*. 2015;359(2):269-74.
101. Currier SC, Lee CK, Chang BS, Bodell AL, Pai GS, Job L, et al. Mutations in POMT1 are found in a minority of patients with Walker-Warburg syndrome. *Am J Med Genet A*. 2005;133A(1):53-7.
102. Bouchet C, Gonzales M, Vuillaumier-Barrot S, Devisme L, Lebizec C, Alanio E, et al. Molecular heterogeneity in fetal forms of type II lissencephaly. *Hum Mutat*. 2007;28(10):1020-7.
103. Messina S, Mora M, Pegoraro E, Pini A, Mongini T, D'Amico A, et al. POMT1 and POMT2 mutations in CMD patients: a multicentric Italian study. *Neuromuscul Disord*. 2008;18(7):565-71.
104. Denny JC, Crawford DC, Ritchie MD, Bielinski SJ, Basford MA, Bradford Y, et al. Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: using electronic medical records for genome- and phenome-wide studies. *Am J Hum Genet*. 2011;89(4):529-42.
105. Wang F, Zang Y, Li M, Liu W, Wang Y, Yu X, et al. *DUOX2* and *DUOXA2* Variants Confer Susceptibility to Thyroid Dysgenesis and Gland-*in-situ* With Congenital Hypothyroidism. *Front Endocrinol (Lausanne)*. 2020;11:237.
106. Jin Y, Wang J, Bachtiar M, Chong SS, Lee CGL. Architecture of polymorphisms in the human genome reveals functionally important and positively selected variants in immune response and drug transporter genes. *Hum Genomics*. 2018;12(1):43.

107. Camargo RY, Kanamura CT, Friguglietti CU, Nogueira CR, Iorcansky S, Tincani AJ, et al. Histopathological Characterization and Whole Exome Sequencing of Ectopic Thyroid: Fetal Architecture in a Functional Ectopic Gland from Adult Patient. *Int J Endocrinol*. 2018;2018:4682876.
108. Peters C, van Trotsenburg ASP, Schoenmakers N. DIAGNOSIS OF ENDOCRINE DISEASE: Congenital hypothyroidism: update and perspectives. *Eur J Endocrinol*. 2018;179(6):R297-R317.
109. Castanet M, Lyonnet S, Bonaïti-Pellié C, Polak M, Czernichow P, Léger J. Familial forms of thyroid dysgenesis among infants with congenital hypothyroidism. *N Engl J Med*. 2000;343(6):441-2.
110. Magne F, Serpa R, Van Vliet G, Samuels ME, Deladoëy J. Somatic mutations are not observed by exome sequencing of lymphocyte DNA from monozygotic twins discordant for congenital hypothyroidism due to thyroid dysgenesis. *Horm Res Paediatr*. 2015;83(2):79-85.
111. Jerome LA, Papaioannou VE. DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. *Nat Genet*. 2001;27(3):286-91.
112. Vitelli F, Taddei I, Morishima M, Meyers EN, Lindsay EA, Baldini A. A genetic link between *Tbx1* and fibroblast growth factor signaling. *Development*. 2002;129(19):4605-11.
113. Aggarwal VS, Liao J, Bondarev A, Schimmang T, Lewandoski M, Locker J, et al. Dissection of *Tbx1* and *Fgf* interactions in mouse models of 22q11DS suggests functional redundancy. *Hum Mol Genet*. 2006;15(21):3219-28.
114. Abu-Khudir R, Paquette J, Lefort A, Libert F, Chanoine JP, Vassart G, et al. Transcriptome, methylome and genomic variations analysis of ectopic thyroid glands. *PLoS One*. 2010;5(10):e13420.
115. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. *Cell*. 2018;172(4):650-65.
116. Clifton-Bligh RJ, Wentworth JM, Heinz P, Crisp MS, John R, Lazarus JH, et al. Mutation of the gene encoding human TTF-2 associated with thyroid agenesis, cleft palate and choanal atresia. *Nat Genet*. 1998;19(4):399-401.
117. Castanet M, Park SM, Smith A, Bost M, Léger J, Lyonnet S, et al. A novel loss-of-function mutation in TTF-2 is associated with congenital hypothyroidism, thyroid agenesis and cleft palate. *Hum Mol Genet*. 2002;11(17):2051-9.
118. Carré A, Hamza RT, Kariyawasam D, Guillot L, Teissier R, Tron E, et al. A novel FOXE1 mutation (R73S) in Bamforth-Lazarus syndrome causing increased thyroidal gene expression. *Thyroid*. 2014;24(4):649-54.
119. Baris I, Arisoy AE, Smith A, Agostini M, Mitchell CS, Park SM, et al. A novel missense mutation in human TTF-2 (FKHL15) gene associated with congenital hypothyroidism but not athyreosis. *J Clin Endocrinol Metab*. 2006;91(10):4183-7.

120. Kang IN, Musa M, Harun F, Junit SM. Characterization of mutations in the FOXE1 gene in a cohort of unrelated Malaysian patients with congenital hypothyroidism and thyroid dysgenesis. *Biochem Genet.* 2010;48(1-2):141-51.
121. Castanet M, Polak M. Spectrum of Human Foxe1/TTF2 Mutations. *Horm Res Paediatr.* 2010;73(6):423-9.
122. Pereira JS, da Silva JG, Tomaz RA, Pinto AE, Bugalho MJ, Leite V, et al. Identification of a novel germline FOXE1 variant in patients with familial non-medullary thyroid carcinoma (FNMTc). *Endocrine.* 2015;49(1):204-14.
123. Landa I, Ruiz-Llorente S, Montero-Conde C, Inglada-Pérez L, Schiavi F, Leskelä S, et al. The variant rs1867277 in FOXE1 gene confers thyroid cancer susceptibility through the recruitment of USF1/USF2 transcription factors. *PLoS Genet.* 2009;5(9):e1000637.
124. Xie L, Deng Y, Yuan Y, Tan X, Liu L, Li N, et al. Association of SNP rs1867277 in FOXE1 Gene and Cleft Lip with or without Cleft Palate in a Han Chinese Population. *Fetal Pediatr Pathol.* 2018;37(2):89-94.
125. Yin X, Zhang H, Zhu Z, Wang H, Du Y, Li S, et al. FOXE1 polymorphisms and non-syndromic orofacial cleft susceptibility in a Chinese Han population. *Oral Dis.* 2016;22(4):274-9.
126. Carré A, Castanet M, Sura-Trueba S, Szinnai G, Van Vliet G, Trochet D, et al. Polymorphic length of FOXE1 alanine stretch: evidence for genetic susceptibility to thyroid dysgenesis. *Hum Genet.* 2007;122(5):467-76.
127. Pimentel CP, Cortinhas-Alves EA, de Oliveira EHC, Santana-da-Silva LC. Does the Polymorphism in the Length of the Polyalanine Tract of. *J Thyroid Res.* 2017;2017:2793205.
128. Kallel R, Belguith-Maalej S, Akdi A, Mnif M, Charfeddine I, Galofré P, et al. Genetic investigation of FOXE1 polyalanine tract in thyroid diseases: new insight on the role of FOXE1 in thyroid carcinoma. *Cancer Biomark.* 2010;8(1):43-51.
129. Bullock M, Duncan EL, O'Neill C, Tacon L, Sywak M, Sidhu S, et al. Association of FOXE1 polyalanine repeat region with papillary thyroid cancer. *J Clin Endocrinol Metab.* 2012;97(9):E1814-9.
130. Nikitski AV, Rogounovitch TI, Bychkov A, Takahashi M, Yoshiura KI, Mitsutake N, et al. Genotype Analyses in the Japanese and Belarusian Populations Reveal Independent Effects of rs965513 and rs1867277 but Do Not Support the Role of FOXE1 Polyalanine Tract Length in Conferring Risk for Papillary Thyroid Carcinoma. *Thyroid.* 2017;27(2):224-35.
131. Brito LA, Cruz LA, Rocha KM, Barbara LK, Silva CB, Bueno DF, et al. Genetic contribution for non-syndromic cleft lip with or without cleft palate (NS CL/P) in different regions of Brazil and implications for association studies. *Am J Med Genet A.* 2011;155A(7):1581-7.

132. de Aquino SN, Messetti AC, Hoshi R, Borges A, Viena CS, Reis SR, et al. Analysis of susceptibility polymorphisms for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. *Birth Defects Res A Clin Mol Teratol*. 2014;100(1):36-42.
133. Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, et al. FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. *Hum Mol Genet*. 2009;18(24):4879-96.
134. do Rego Borges A, Sá J, Hoshi R, Viena CS, Mariano LC, de Castro Veiga P, et al. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry. *Am J Med Genet A*. 2015;167A(10):2344-9.
135. Tomaz RA, Sousa I, Silva JG, Santos C, Teixeira MR, Leite V, et al. FOXE1 polymorphisms are associated with familial and sporadic nonmedullary thyroid cancer susceptibility. *Clin Endocrinol (Oxf)*. 2012;77(6):926-33.
136. Chrisoulidou A, Boudina M, Tzemailas A, Doumala E, Iliadou PK, Patakiouta F, et al. Histological subtype is the most important determinant of survival in metastatic papillary thyroid cancer. *Thyroid Res*. 2011;4(1):12.
137. Mendieta-Zerón H, Jiménez-Rosales A, Pérez-Amado CJ, Jiménez-Morales S. Mutation Screening in a Case with Cleft Lip, Hypothyroidism, and Thyroid Carcinoma: A New Syndrome? *Case Rep Genet*. 2017;2017:6390545.
138. Bonora E, Rizzato C, Diquigiovanni C, Oudot-Mellakh T, Campa D, Vargiolu M, et al. The FOXE1 locus is a major genetic determinant for familial nonmedullary thyroid carcinoma. *Int J Cancer*. 2014;134(9):2098-107.
139. Frezzetti D, Reale C, Cali G, Nitsch L, Fagman H, Nilsson O, et al. The microRNA-processing enzyme Dicer is essential for thyroid function. *PLoS One*. 2011;6(11):e27648.
140. Bartram MP, Amendola E, Benzing T, Schermer B, de Vita G, Müller RU. Mice lacking microRNAs in Pax8-expressing cells develop hypothyroidism and end-stage renal failure. *BMC Mol Biol*. 2016;17:11.
141. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol*. 2013;14(8):475-88.
142. Ipsaro JJ, Joshua-Tor L. From guide to target: molecular insights into eukaryotic RNA-interference machinery. *Nat Struct Mol Biol*. 2015;22(1):20-8.
143. Robinson WP. Mechanisms leading to uniparental disomy and their clinical consequences. *Bioessays*. 2000;22(5):452-9.
144. McLean R, Howard N, Murray IP. Thyroid dysgenesis in monozygotic twins: variants identified by scintigraphy. *Eur J Nucl Med*. 1985;10(7-8):346-8.

145. Westerlund J, Andersson L, Carlsson T, Zoppoli P, Fagman H, Nilsson M. Expression of *Islet1* in thyroid development related to budding, migration, and fusion of primordia. *Dev Dyn*. 2008;237(12):3820-9.
146. Vandernoot I, Haerlingen B, Gillotay P, Trubiroha A, Janssens V, Opitz R, et al. Enhanced Canonical Wnt Signaling During Early Zebrafish Development Perturbs the Interaction of Cardiac Mesoderm and Pharyngeal Endoderm and Causes Thyroid Specification Defects. *Thyroid*. 2021;31(3):420-38.
147. Marelli F, Persani L. Role of Jagged1-Notch pathway in thyroid development. *J Endocrinol Invest*. 2018;41(1):75-81.
148. Long W, Guo F, Yao R, Wang Y, Wang H, Yu B, et al. Genetic and Phenotypic Characteristics of Congenital Hypothyroidism in a Chinese Cohort. *Front Endocrinol (Lausanne)*. 2021;12:705773.
149. Makretskaya N, Bezlepina O, Kolodkina A, Kiyayev A, Vasilyev EV, Petrov V, et al. High frequency of mutations in 'dyshormonogenesis genes' in severe congenital hypothyroidism. *PLoS One*. 2018;13(9):e0204323.
150. Sage EH, Reed M, Funk SE, Truong T, Steadele M, Puolakkainen P, et al. Cleavage of the matricellular protein SPARC by matrix metalloproteinase 3 produces polypeptides that influence angiogenesis. *J Biol Chem*. 2003;278(39):37849-57.
151. Minna E, Brich S, Todoerti K, Pilotti S, Collini P, Bonaldi E, et al. Cancer Associated Fibroblasts and Senescent Thyroid Cells in the Invasive Front of Thyroid Carcinoma. *Cancers (Basel)*. 2020;12(1).
152. Kopp P. Thyroid hormone synthesis. In: Braverman L, Cooper D, editors. *Werner & Ingbar's The Thyroid: A Fundamental and Clinical Text*. 10th. Philadelphia: Lippincott Williams & Wilkins; 2013. p. 48-74.
153. Song Y, Ruf J, Lothaire P, Dequanter D, Andry G, Willemse E, et al. Association of duoxes with thyroid peroxidase and its regulation in thyrocytes. *J Clin Endocrinol Metab*. 2010;95(1):375-82.
154. Van Vliet G. Development of the thyroid gland: lessons from congenitally hypothyroid mice and men. *Clin Genet*. 2003;63(6):445-55.
155. Mahjabeen I, Baig RM, Sabir M, Kayani MA. Genetic and expressional variations of *APEX1* are associated with increased risk of head and neck cancer. *Mutagenesis*. 2013;28(2):213-8.
156. Pocar P, Klonisch T, Brandsch C, Eder K, Fröhlich C, Hoang-Vu C, et al. AhR-agonist-induced transcriptional changes of genes involved in thyroid function in primary porcine thyrocytes. *Toxicol Sci*. 2006;89(2):408-14.
157. Opitz R, Maquet E, Zoenen M, Dadhich R, Costagliola S. TSH receptor function is required for normal thyroid differentiation in zebrafish. *Mol Endocrinol*. 2011;25(9):1579-99.

158. Catalano V, Dentice M, Ambrosio R, Luongo C, Carollo R, Benfante A, et al. Activated Thyroid Hormone Promotes Differentiation and Chemotherapeutic Sensitization of Colorectal Cancer Stem Cells by Regulating Wnt and BMP4 Signaling. *Cancer Res.* 2016;76(5):1237-44.
159. Panicker V. Genetics of thyroid function and disease. *Clin Biochem Rev.* 2011;32(4):165-75.
160. Ross AW, Helfer G, Russell L, Darras VM, Morgan PJ. Thyroid hormone signalling genes are regulated by photoperiod in the hypothalamus of F344 rats. *PLoS One.* 2011;6(6):e21351.
161. Leone V, D'Angelo D, Ferraro A, Pallante P, Rubio I, Santoro M, et al. A TSH-CREB1-microRNA loop is required for thyroid cell growth. *Mol Endocrinol.* 2011;25(10):1819-30.
162. Kurihara Y, Kurihara H, Maemura K, Kuwaki T, Kumada M, Yazaki Y. Impaired development of the thyroid and thymus in endothelin-1 knockout mice. *J Cardiovasc Pharmacol.* 1995;26 Suppl 3:S13-6.
163. Elsalini OA, von Gartzen J, Cramer M, Rohr KB. Zebrafish *hhex*, *nk2.1a*, and *pax2.1* regulate thyroid growth and differentiation downstream of Nodal-dependent transcription factors. *Dev Biol.* 2003;263(1):67-80.
164. Cooley MA, Kern CB, Fresco VM, Wessels A, Thompson RP, McQuinn TC, et al. Fibulin-1 is required for morphogenesis of neural crest-derived structures. *Dev Biol.* 2008;319(2):336-45.
165. Hannenhalli S, Kaestner KH. The evolution of Fox genes and their role in development and disease. *Nat Rev Genet.* 2009;10(4):233-40.
166. Zhu P, Piao Y, Dong X, Jin Z. Forkhead box J1 expression is upregulated and correlated with prognosis in patients with clear cell renal cell carcinoma. *Oncol Lett.* 2015;10(3):1487-94.
167. Kameda Y, Ito M, Nishimaki T, Gotoh N. *FRS2a* is required for the separation, migration, and survival of pharyngeal endoderm derived organs including thyroid, ultimobranchial body, parathyroid and thymus. In: Carel J-C, Rochberg Ze, editors. *Yearbook of Pediatric Endocrinology 2009*. Basel, Switzerland: Karger; 2009. p. 32-3.
168. Ferlin A, De Toni L, Agoulnik AI, Lunardon G, Armani A, Bortolanza S, et al. Protective Role of Testicular Hormone INSL3 From Atrophy and Weakness in Skeletal Muscle. *Front Endocrinol (Lausanne).* 2018;9:562.
169. Vieira JM, Santos SC, Espadinha C, Correia I, Vag T, Casalou C, et al. Expression of vascular endothelial growth factor (VEGF) and its receptors in thyroid carcinomas of follicular origin: a potential autocrine loop. *Eur J Endocrinol.* 2005;153(5):701-9.
170. Kizys MM, Nesi-França S, Cardoso MG, Harada MY, Melo MC, Chiamolera MI, et al. The absence of mutations in homeobox candidate genes *HOXA3*, *HOXB3*,

HOXD3 and PITX2 in familial and sporadic thyroid hemiagenesis. *J Pediatr Endocrinol Metab.* 2014;27(3-4):317-22.

171. Netchine I, Sobrier ML, Krude H, Schnabel D, Maghnie M, Marcos E, et al. Mutations in LHX3 result in a new syndrome revealed by combined pituitary hormone deficiency. *Nat Genet.* 2000;25(2):182-6.

172. Consortium G. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369(6509):1318-30.

173. Enomoto K, Zhu X, Park S, Zhao L, Zhu YJ, Willingham MC, et al. Targeting MYC as a Therapeutic Intervention for Anaplastic Thyroid Cancer. *J Clin Endocrinol Metab.* 2017;102(7):2268-80.

174. Porazzi P, Calebiro D, Benato F, Tiso N, Persani L. Thyroid gland development and function in the zebrafish model. *Mol Cell Endocrinol.* 2009;312(1-2):14-23.

175. Opitz R, Hitz MP, Vandernoot I, Trubiroha A, Abu-Khudir R, Samuels M, et al. Functional zebrafish studies based on human genotyping point to netrin-1 as a link between aberrant cardiovascular development and thyroid dysgenesis. *Endocrinology.* 2015;156(1):377-88.

176. Ding Z, Ke R, Zhang Y, Fan Y, Fan J. FOXE1 inhibits cell proliferation, migration and invasion of papillary thyroid cancer by regulating PDGFA. *Mol Cell Endocrinol.* 2019;493:110420.

177. Wassner AJ, Cohen LE, Hechter E, Dauber A. Isolated central hypothyroidism in young siblings as a manifestation of PROP1 deficiency: clinical impact of whole exome sequencing. *Horm Res Paediatr.* 2013;79(6):379-86.

178. Ramos HE, Carré A, Chevrier L, Szinnai G, Tron E, Cerqueira TL, et al. Extreme phenotypic variability of thyroid dysgenesis in six new cases of congenital hypothyroidism due to PAX8 gene loss-of-function mutations. *Eur J Endocrinol.* 2014;171(4):499-507.

179. Naiche LA, Harrelson Z, Kelly RG, Papaioannou VE. T-box genes in vertebrate development. *Annu Rev Genet.* 2005;39:219-39.

180. Pisarev MA, Thomasz L, Juvenal GJ. Role of transforming growth factor beta in the regulation of thyroid function and growth. *Thyroid.* 2009;19(8):881-92.

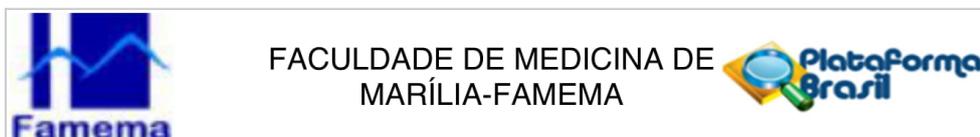
181. Oishi T, Iida A, Otsubo S, Kamatani Y, Usami M, Takei T, et al. A functional SNP in the NKX2.5-binding site of ITPR3 promoter is associated with susceptibility to systemic lupus erythematosus in Japanese population. *J Hum Genet.* 2008;53(2):151-62.

182. Stoupa A, Adam F, Kariyawasam D, Strassel C, Gawade S, Szinnai G, et al. TUBB1 mutations cause thyroid dysgenesis associated with abnormal platelet physiology. *EMBO Mol Med.* 2018;10(12).

183. Xia S, Ji R, Xu Y, Ni X, Dong Y, Zhan W. Twisted Gastrulation BMP Signaling Modulator 1 Regulates Papillary Thyroid Cancer Cell Motility and Proliferation. *J Cancer*. 2017;8(14):2816-27.
184. Leoni SG, Kimura ET, Santisteban P, De la Vieja A. Regulation of thyroid oxidative state by thioredoxin reductase has a crucial role in thyroid responses to iodide excess. *Mol Endocrinol*. 2011;25(11):1924-35.
185. Fernandez LP, Lopez-Marquez A, Martinez AM, Gomez-Lopez G, Santisteban P. New insights into FoxE1 functions: identification of direct FoxE1 targets in thyroid cells. *PloS one*. 2013;8(5):e62849.
186. Di Palma T, Conti A, de Cristofaro T, Scala S, Nitsch L, Zannini M. Identification of novel Pax8 targets in FRTL-5 thyroid cells by gene silencing and expression microarray analysis. *PLoS One*. 2011;6(9):e25162.
187. Ruiz-Llorente S, Carrillo Santa de Pau E, Sastre-Perona A, Montero-Conde C, Gómez-López G, Fagin JA, et al. Genome-wide analysis of Pax8 binding provides new insights into thyroid functions. *BMC Genomics*. 2012;13:147.

Anexos

Anexo 1: Parecer de aprovação do Comitê de Ética em Pesquisa da UNIFESP (continua)



PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Identificação de alterações genéticas e cromossômicas no exoma completo de pacientes com hipotireoidismo congênito, por disgenesia tireoidiana

Pesquisador: Ileana Rubio

Área Temática: Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP;);

Versão: 1

CAAE: 59240816.2.3001.5413

Instituição Proponente: FACULDADE DE MEDICINA DE MARILIA

Patrocinador Principal: FUNDACAO DE AMPARO A PESQUISA DO ESTADO DE SAO PAULO

DADOS DO PARECER

Número do Parecer: 2.916.115

Apresentação do Projeto:

O hipotireoidismo congênito (HC) constitui a síndrome endócrina pediátrica mais frequente cuja principal causa é a disgenesia tireoidiana (DT). A DT decorre de defeitos embriológicos que provocam agenesia glandular, ectopia ou hipoplasia. Até o momento, as causas genéticas foram identificadas em somente 5% dos casos de DT, envolvendo mutações nos genes PAX8 e do receptor de TSH, e em FOXE1, NKX2.1, NKX2.5 e TBX1 associadas a outras malformações. Neste estudo propomos ampliar a investigação das bases moleculares da disgenesia tireoidiana através da pesquisa de alterações gênicas no exoma completo e regiões regulatórias (UTRs e de ligação de microRNAs) de pacientes com DT, visando identificar aquelas com significado clínico, através de sequenciamento de nova geração. Para tanto, será realizada a construção de biblioteca e captura que permitirá sequenciar 71Mb por amostra (SureSelect Exome V6+UTR Capture Library, Agilent) em plataforma Illumina® NextSeq™ 500.

Objetivo da Pesquisa:

Objetivo Primário: Este projeto tem como objetivo ampliar o estudo das bases moleculares do hipotireoidismo congênito causado por disgenesia tireoidiana, através da pesquisa de alterações gênicas com relevância clínica no exoma completo e em sequências regulatórias (regiões UTRs e de ligação de microRNAs), utilizando sequenciamento de nova geração, em uma coorte de

Endereço: Av: Monte Carmelo, 800 - Sala 04

Bairro: Fragata

CEP: 17.519-030

UF: SP

Município: MARILIA

Telefone: (14)3402-1744

Fax: (14)3422-1079

E-mail: dirpos@famema.br

Anexo 1: Parecer de aprovação do Comitê de Ética em Pesquisa da UNIFESP (continuação)



Continuação do Parecer: 2.916.115

pacientes com DT cuidadosamente caracterizada em termos bioquímicos e morfológicos.

Avaliação dos Riscos e Benefícios:

Riscos: Não há riscos.

Benefícios: Este estudo é inédito e contribuirá na literatura com o conhecimento das bases moleculares do HC e das disgenesia tireoidiana e permitirá aprofundar o conhecimento do desenvolvimento "normal" da glândula tireóide.

Comentários e Considerações sobre a Pesquisa:

Análises de bioinformática Os resultados do sequenciamento serão analisados considerando 1) a casuística toda, para investigar alterações associadas à DT e 2) cada um dos grupos: ectopia, atireose ou hipoplasia, para avaliar características genéticas específicas de cada fenótipo. As sequências serão avaliadas quanto à qualidade no software R (v.3.1.2) por meio do pacote Rqc, disponível no repositório Bioconductor, e alinhadas com sequência de referência do genoma humano (GRCh37/hg19) por meio do alinhador BWA-MEM e indexados no samtools. As sequências serão analisadas usando ferramentas específicas do software GATK, as variantes identificadas serão anotadas usando os softwares snpEff e ANNOVAR, de forma a atribuir no conhecimento biológico a respeito de sua função. As anotações escolhidas incluirão dados sobre a região genômica e genes, frequência populacional (1000 Genomes, ESP6500), impacto funcional, escores de predição de patogenicidade (como pelo PolyPhen-2), conservação, entre outros. Serão utilizados também algoritmos específicos para identificar variações de número de cópias por meio das ferramentas ExomeCNV, Conifer, Canoes e outras aberrações cromossômicas como dissomia uniparental. Após a identificação das variantes (SNPs, indels e CNVs) e sua anotação funcional, estas serão priorizadas para posterior análise e validação por meio da aplicação de filtros que levam em consideração o efeito funcional na proteína, a frequência populacional em bancos de dados como o 1000 Genomes e ESP6500, além de menção em bancos de dados de alterações genômicas estruturais (no caso das CNVs), como DGV (Database of Genomic Variants) e DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources) ou de mutações e variantes pontuais como o Clinvar. O software DAVID será utilizado para classificação funcional de genes com variantes de interesse, por meio de classificação com Gene Ontology, agrupamento de famílias gênicas, e vias de sinalização. Análises para montagem de redes gênicas também serão

Endereço: Av: Monte Carmelo, 800 - Sala 04
Bairro: Fragata **CEP:** 17.519-030
UF: SP **Município:** MARILIA
Telefone: (14)3402-1744 **Fax:** (14)3422-1079 **E-mail:** dirpos@famema.br

Anexo 1: Parecer de aprovação do Comitê de Ética em Pesquisa da UNIFESP (continuação)



Continuação do Parecer: 2.916.115

realizadas os genes identificados com o auxílio das ferramentas GeneMANIA e Cytoscape. Validação da variantes gênicas (SNPs e pequenas indels)As alterações identificadas serão validadas através de sequenciamento de Sanger. As regiões serão amplificadas por PCR, utilizando iniciadores específicos desenhados com o auxílio do programa PRIMER 3. O produto de PCR será purificado com enzima ExoSAPIT® (USB Corporation, EUA) e sequenciado utilizando o estojo BigDye® Terminator v3.1 (Life Technologies, Foster City, CA) em equipamento ABI PRISM 3130xl (Life Technologies, Foster City, CA). O resultado será analisado e comparado à sequência referência (GRCh37/hg19). Validação de CNVs por PCR quantitativa (qPCR)O DNA extraído será utilizado para a reação de qPCR utilizando o estojo SYBR® Green Master Mix (Thermo Fisher Scientific, EUA). Os primers para cada CNV identificada serão localizados na região de CNV sobreposta entre as amostras de um mesmo grupo. Os primers serão desenhados utilizando o PRIMER 3, com o auxílio das ferramentas DNA Calculator e IDT OligoAnalyzer. Como controle endógeno serão desenhados primers próximos, mas fora da região de CNV.O ensaios de qPCR serão submetidos à ciclagem no aparelho Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, EUA). As reações serão realizadas em triplicata e o valor de Ct (threshold cycle) será obtido utilizando o Applied Biosystem software (Applied Biosystems, Foster City, CA, EUA), as triplicatas serão submetidas à análise do desvio padrão ($SD < 1$).O número de cópias das regiões encontradas será avaliado por quantificação relativa, pelo método comparativo de Ct (2^{-Ct}), onde $Ct = Ct \text{ amostra teste} - Ct \text{ controle referência}$). Análise estatísticaAs análises estatísticas serão realizadas utilizando o programa GraphPad Prism 5.01 utilizando cálculos adequados para os resultado obtidos, sendo um valor de $p < 0,05$ considerado estatisticamente significativo

Neste estudo serão incluídos 45 pacientes com HC por DT agrupados por fenótipo: ectopia, atireose e hipoplasia ($n=15$ por grupo).

Todos casos esporádicos, dos quais dispomos de DNA de sangue periférico, sendo que os pacientes com ectopia e agenesia não apresentam mutações no gene PAX8 e os hipoplásicos no TSHR. O diagnóstico neonatal e o acompanhamento desses pacientes está sendo realizado pela APAE Salvador, em colaboração com o Dr Prof Dr Helton Ramos da Universidade Federal da Bahia, UFBA.

Endereço: Av: Monte Carmelo, 800 - Sala 04
Bairro: Fragata **CEP:** 17.519-030
UF: SP **Município:** MARILIA
Telefone: (14)3402-1744 **Fax:** (14)3422-1079 **E-mail:** dirpos@famema.br

Anexo 1: Parecer de aprovação do Comitê de Ética em Pesquisa da UNIFESP (continuação)



Continuação do Parecer: 2.916.115

Amostras de DNA de sangue periférico de indivíduos sem doença tireoidiana serão utilizadas como controles para responder se as variantes podem ser consideradas mutações ou polimorfismos. Já dispomos de 150 amostras de DNA de indivíduos controles normais.

Para o sequenciamento global do exoma expandido (WES-expandido), será empregada a plataforma Illumina® NextSeq™ 500 (Illumina, Inc., EUA), em parceria com o Centro de Oncologia Molecular, Instituto Sírio-Libanês de ensino e Pesquisa. A primeira etapa consiste na construção de bibliotecas e o enriquecimento do exoma utilizando o SureSelectXT Human All Exon V6+UTR Capture Library (Agilent Technologies, EUA), que enriquece o exoma, regiões UTR, regiões de ligação de microRNAs, capturando 71 Mb por amostra. Para o sequenciamento de alta performance será utilizado o kit NextSeq 500 High Output v2 (300Cycles) (Illumina Inc. EUA) com cobertura média calculada e esperada para cada amostra de aproximadamente 140X.

Considerações sobre os Termos de apresentação obrigatória:

A folha de rosto esta preenchida, datada e assinada pelo pesquisador principal e pelo responsável institucional.

Os pesquisadores se comprometem com o sigilo e confidencialidade de todo e qualquer dado obtido, bem como com a privacidade dos participantes da pesquisa.

TCLE:

As amostras de sangue utilizadas neste estudo foram utilizadas em estudo anterior intitulado "Pesquisa de mutações de genes implicados na morfogênese tireoidiana e cardiovascular: estudo clínico e molecular em coortes de pacientes com disgenesia tireoidiana e cardiopatia congênita" que foi aprovado pela comissão de Ética Médica do Hospital Santa Izabel (Parecer N° 47/11) e Hospital Universitário Edgar Santos (Parecer N° 125/2011) e todos os participantes deram consentimento formal antes da coleta e conservação de material biológico, após informação sobre o estudo.

Recomendações:

Nenhuma

Conclusões ou Pendências e Lista de Inadequações:

Aprovado

Considerações Finais a critério do CEP:

Diante do exposto, o CEP FAMEMA, de acordo com as atribuições definidas na Resolução CNS

Endereço: Av: Monte Carmelo, 800 - Sala 04
Bairro: Fragata **CEP:** 17.519-030
UF: SP **Município:** MARILIA
Telefone: (14)3402-1744 **Fax:** (14)3422-1079 **E-mail:** dirpos@famema.br

Anexo 1: Parecer de aprovação do Comitê de Ética em Pesquisa da UNIFESP (conclusão)



Continuação do Parecer: 2.916.115

466/2012 e na Norma Operacional N° 001/2013 do CNS manifesta-se pela Aprovação do Projeto de Pesquisa.

Aprovado: Retirar Documentos assinados pelo CEP/FAMEMA após 01/10/18

Observação: O CEP FAMEMA informa que, a partir da data de aprovação, é necessário o envio de relatórios parciais (anualmente), e o relatório final, quando do término do estudo

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Outros	Justificativa_emenda_FAMEMA.pdf	17/08/2018 11:22:21	Ileana Rubio	Aceito
Outros	Sgarbiparticipante.pdf	10/07/2018 11:04:02	Ileana Rubio	Aceito
Outros	Justificativa_da_emenda2.doc	27/02/2018 09:59:08	Ileana Rubio	Aceito
Projeto Detalhado / Brochura Investigador	Projeto.pdf	20/04/2016 11:12:23	Ileana Rubio	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	20/04/2016 11:08:57	Ileana Rubio	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

MARILIA, 25 de Setembro de 2018

Assinado por:
Maria José Sanches Marin
(Coordenador(a))

Endereço: Av: Monte Carmelo, 800 - Sala 04
Bairro: Fragata **CEP:** 17.519-030
UF: SP **Município:** MARILIA
Telefone: (14)3402-1744 **Fax:** (14)3422-1079 **E-mail:** dirpos@famema.br

Anexo 2: Dados clínicos e laboratoriais dos 35 pacientes selecionados para sequenciamento WES

Pt	Gn	TSH neo (mUI/mL)	TSH (mUI/mL)	T4L (ng/dL)	T4T (ng/dL)	Tg (ng/mL)	USG	Fenótipo
1	F	182	100	0,04	0,53	0,2	AG	AG
2	F	344	100	0,08	0,8	0,2	AG	AG
3	M	124	252	0,4	-	0,5	AG	AG
4	F	468	100	0,02	0,42	0,5	AG	AG
5	F	1057,6	100	0,808	7,68	0,8	AG	AG
6	F	701	-	-	-	0,2	AG	AG
7	F	1397,5	27,56	2,46	18,75	1	AG	AG
8	F	150	-	-	-	1	AG	AG
9	F	-	-	-	-	0,2	AG	AG
10	F	100	17,49	1,07	17,49	0,2	AG	AG
11	F	100	48,83	1,78	14,91	1,2	TP	HP
12	M	332	100	0,18	1,47	2,1	TP	HP
13	M	36	158,3	-	4,54	10	TP	HP
14	F	174,1	100	0,588	5,33	23,6	TP	HP
15	M	124	100	0,131	1,5	3,4	TP	HP
16	F	12,9	83,17	0,431	3,5	7,8	TP	HP
17	F	41	35,88	1,04	9,77	12,9	TP	HP
18	M	24,9	100	0,46	3,34	18,2	TP	HP
19	M	402	4,21	1,04	8,84	6	TP	HP
20	M	10,3	10,53	1,33	-	23,5	TP	HP
21	F	40,7	19,1	0,7	6,03	5,7	TP	HP
22	M	-	100	0,22	4,43	6,5	TP	HP
23	F	-	-	-	-	35,8	TP	HP
24	F	562	222,35	0,01	0,1	128	AG	EC
25	F	53,6	86,72	0,81	6	16,9	AG	EC
26	F	91,4	100	-	-	4,5	AG	EC
27	F	19,3	34,59	1,08	6,63	1,9	AG	EC
28	F	12,8	0,136	2,61	21,6	47,4	AG	EC
29	M	48,9	43,47	0,81	6,65	2,1	AG	EC
30	F	91,5	95	0,584	4,82	16	AG	EC
31	M	174	100	0,32	1,05	17,3	AG	EC
32	M	358	100	-	1,53	28,7	AG	EC
33	F	44,8	14,63	1,2	11,95	73,2	AG	EC
34	F	-	-	-	-	-	AG	EC
35	M	-	-	-	-	-	AG	EC
Valor de referência		< 10 mU/L	0,3-4,0 mU/L	0,9-2,6 µg/dL	8,2-16,6 µg/dL	2,0-70 ng/mL		

Pt: paciente; Gn: gênero; USG: ultrassonografia; TP: tópica; AG: agenesia; HP: hipoplasia; EC: ectopia.

Anexo 3: *Primers* (iniciadores) utilizados nas análises de sequenciamento *Sanger*

Gene	Sequência do <i>primer forward</i> (5' - 3')	Sequência do <i>primer reverse</i> (5' - 3')
<i>DUOX2</i> _rs764517475	GGGAAGCTGAAGTCCTGAGA	TGAAGATTTGGCCTCTGTCG
<i>DUOX2</i> _rs200459845	GCTGGAGCCCCTCCT	GACCCCATGGCCAGTATAGA
<i>DUOX2</i> _rs113632824	AGTCTGTCCTGGTTGGCATC	CACCCTATGAGTCCCAGGAG
<i>DUOX2</i> _rs530719719	GTTCTCCTGGCTGCAAAGAC	GATATGTGGGTGGGGCCTA
<i>DUOX2</i> _rs143818717	CTTCCCATCCCAGTACTTC	GCACCCTCAATCTTGATCCT
<i>DUOX2</i> _rs201197899	GAGGGATGGGGCAACAGT	AGGCTGAGGAGCAGTCTGAG
<i>FOXE1</i> _rs907577	GCTACCTCTACCACACAGTT	TGAAAGGACGGGCTCAT
<i>FOXE1</i> _rs1867277	GCCCAGCGCCAGTACTAACT	GGCTACCGTGAAGGAAGAGC
<i>FOXE1</i> _rs71369530	CGGGTAAGGGCAACTACTG	AGGAACCAGGCCGAAGA
<i>TSHR</i> _rs71369530	TGGACCGGAAGATCCGCCTC	CGAGACCGTATATAACAAA

Anexo 4: Dados de sequenciamento das 35 amostras de disgenesias selecionadas para WES

Paciente	Total de bases	Bases não cobertas	Bases cobertas (%)					
			1-30x	≥30x	≥ 50x	≥100x	≥150x	≥200x
1	85462805	813	35,88	64,12	35,13	8,87	3,18	0,91
2	85497073	720	32,94	67,06	38,98	11,33	4,34	1,42
3	85523735	751	29,27	70,73	43,67	11,9	3,84	1,12
4	85487248	714	29,93	70,07	43,16	11,56	3,97	1,27
5	85534264	734	28,41	71,59	44,57	12,28	4,12	1,28
6	85586552	555	22,16	77,84	54,14	17,15	7,22	3,33
7	85533572	771	24,04	75,96	51,88	16,16	7,29	3,62
8	85587680	675	19,37	80,63	59,88	21,55	9,92	5,55
9	85560843	711	23,06	76,94	53,55	17,26	7,82	3,98
10	85544084	813	25,44	74,56	49,44	15,05	6,51	2,95
11	85442290	810	35,48	64,52	36,13	10,01	4,06	1,4
12	85538689	636	22,57	77,43	54,21	17,37	7,53	3,65
13	85581752	566	18,83	81,17	59,24	20,67	8,42	3,81
14	85537335	634	25,16	74,84	49,24	14,49	5,09	1,74
15	85557368	575	20,75	79,25	56,03	18,49	7,17	2,96
16	85574650	583	19,56	80,44	57,67	19,81	7,78	3,22
17	85587560	652	20,44	79,56	56,26	18,89	7,55	3,23
18	85456008	798	35,82	64,18	35,28	8,83	2,99	0,84
19	85274426	509	14,89	85,11	66,69	27,59	12,08	6,17
20	85604253	537	16,37	83,63	63,16	23,93	10,22	5,02
21	85490385	837	36,80	63,20	34,16	8,29	2,45	0,56
22	85564166	674	23,45	76,55	51,61	15,89	6,48	2,77
23	85506296	845	27,05	72,95	47,72	14,18	6,44	3,10
24	85510700	713	32,09	67,91	39,26	10,03	3,56	1,05
25	85469238	870	37,11	62,89	33,67	7,99	2,39	0,54
26	85534821	622	23,87	76,13	51,1	15,41	5,74	2,16
27	85559233	600	23,58	76,42	51,58	15,66	5,96	2,32
28	85496000	752	29,54	70,46	43,69	11,7	3,99	1,29
29	85540247	689	24,44	75,56	50,25	14,84	5,63	2,16
30	85529796	728	27,33	72,67	46,44	13,39	5,26	2,05
31	85564594	647	21,42	78,58	54,78	17,76	7,37	3,32
32	85577015	633	19,18	80,82	58,82	20,45	8,7	4,19
33	85584636	643	19,4	80,6	58,27	20,2	8,44	3,91
34	85450281	818	35,4	64,59	36,18	9,05	3,02	0,85
35	85570230	701	22	78	54,93	18,67	7,53	3,24

Anexo 5: Genes candidatos avaliados neste estudo (continua)

Gene	Nome	Ref
ABCA13	ATP binding cassette subfamily A member 13	(114)
ACTA1	Actin, alpha 1, skeletal muscle	(114)
ANKRD36B	Ankyrin repeat domain 36B	(107)
APEX1	Apurinic/Apyrimidinic Endodeoxyribonuclease 1	(155)
ARNT	Aryl Hydrocarbon Receptor Nuclear Translocator	(156)
ATP2A1	ATPase Sarcoplasmic/Endoplasmic Reticulum Ca ²⁺ Transporting 1	(114)
BCL2L1	BCL2 Like 1	(157)
BCL2L12	BCL2 like 12	(157)
BGN	Biglycan	(114)
BMP4	Bone morphogenetic protein 4	(158)
C9orf70	GLIS3 antisense RNA 1	(114)
CCDC114	Coiled-coil domain containing 114	(53)
CDCA8	Cell Division Cycle Associated 8	(68)
CECR1	Cat eye syndrome chromosome region, candidate 1	(114)
CFC1B	Cripto, FRL-1, Cryptic Family 1B	(159)
CGA	Glycoprotein Hormones, Alpha Polypeptide	(160)
CHGA	Chromogranin A	(114)
CHRD	Chordin	(4)
CKM	Creatine kinase, M-type	(114)
COL1A1	Collagen type I alpha 1 chain	(114)
COL3A1	Collagen type III alpha 1 chain	(114)
CPEB4	Cytoplasmic polyadenylation element binding protein 4	(114)
CREB1	cAMP responsive element binding protein 1	(161)
CYBB	Cytochrome b-245 beta chain	(114)
DUOX1	Dual oxidase 1	(6)
DUOX2	Dual oxidase 2	(6)
DUOXA1	Dual oxidase maturation factor 1	(6)
DUOXA2	Dual oxidase maturation factor 2	(6)
DUSP6	Dual specificity phosphatase 6	(114)
EDN1	Endothelin 1	(162)
EDN3	Endothelin 3	(114)
EEF1A2	Eukaryotic translation elongation factor 1 alpha 2	(114)
EGR1	Early growth response 1	(114)
ENO3	Enolase 3	(114)
EYA1	EYA Transcriptional Coactivator And Phosphatase 1	(6)
FAU	Ubiquitin Like And Ribosomal Protein S30 Fusion	(163)
FBLN1	Fibulin 1	(164)
FGF10	Fibroblast growth factor 10	(113)

Anexo 5: Genes candidatos avaliados neste estudo (continuação)

Gene	Nome	Ref
FGF12	Fibroblast growth factor 12	(113)
FGF2	Fibroblast growth factor 2	(113)
FGF3	Fibroblast Growth Factor 3	(113)
FGF8	Fibroblast growth factor 8	(113)
FGFR1	Fibroblast growth factor receptor 1	(113)
FGFR2	Fibroblast growth factor receptor 2	(113)
FGL2	Fibrinogen like 2	(114)
FLJ11127	Family with sequence similarity 105 member A	(114)
FLJ32115	Endoplasmic reticulum protein 27	(114)
FOS	Fos proto-oncogene, AP-1 transcription factor subunit	(114)
FOSB	FosB proto-oncogene, AP-1 transcription factor subunit	(114)
FOXA1	Forkhead box A1	(165)
FOXA2	Forkhead box A2	(165)
FOXE1	Forkhead box E1	(165)
FOXE3	Forkhead Box E3	(165)
FOXJ1	Forkhead box J1	(166)
FOXQ1	Forkhead Box Q1	(165)
FRS2	Fibroblast growth factor receptor substrate 2	(167)
FRS2A	Fibroblast Growth Factor Receptor Substrate 2	(167)
FRZB	Frizzled-related protein	(114)
GATA5	GATA Binding Protein 5	(163)
GLI2	GLI Family Zinc Finger 2	(6)
GLIS3	GLIS family zinc finger 3	(114)
GPNMB	Glycoprotein nmb	(114)
GSTT1	Glutathione S-transferase theta 1	(114)
HADHA	Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Subunit Alpha	(114)
HAND2	Heart- and neural crest derivatives-expressed protein 2	(12)
HES1	Hes Family BHLH Transcription Factor 1	(6)
HESX1	HESX homeobox 1	(6)
HHEX	Hematopoietically-expressed homeobox	(47)
HLA-DQA1	Major histocompatibility complex, class II, DQ alpha 1	(114)
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1	(114)
HOXA2	Homeobox A2	(6)
HOXA3	Homeobox A3	(6)
HOXA5	Homeobox A5	(6)
HOXB3	Homeobox B3	(6)
HOXD3	Homeobox D3	(6)
IGHG4	Immunoglobulin heavy constant gamma 4 (G4m marker)	(114)

Anexo 5: Genes candidatos avaliados neste estudo (continuação)

Gene	Nome	Ref
IGJ	ISL LIM homeobox 2	(114)
IGSF1	Immunoglobulin superfamily member 1	(114)
INSL3	Insulin Like 3	(168)
ISL1	ISL LIM Homeobox 1	(6)
ISL2	ISL LIM homeobox 2	(6)
IYD	Iodotyrosine deiodinase	(6)
JAG1	Jagged 1	(159)
JAG3	Jagged 3	(159)
KDR	Kinase Insert Domain Receptor	(169)
KLF4	Kruppel like factor 4	(114)
LEFTY1	Left-Right Determination Factor 1	(170)
LHX3	LIM homeobox 3	(171)
LHX4	LIM homeobox 4	(171)
LOC440696	Homo sapiens similar to Heat shock protein HSP 90-alpha	(114)
LYZ	Lysozyme	(114)
MIXL1	Mix Paired-Like Homeobox	(172)
MKI67	Marker of proliferation Ki-67	(107)
MKRN1	Makorin ring finger protein 1	(114)
MS4A6A	Membrane spanning 4-domains A6A	(114)
MYBPC1	Myosin binding protein C, slow type	(114)
MYC	MYC Proto-Oncogene, BHLH Transcription Factor	(173)
NDR2	NDRG Family Member 2	(174)
NKX2-1	NK2 homeobox 1	(6)
NKX2-5	NK2 homeobox 5	(6)
NTN1	Netrin 1	(175)
PABPC1	Poly(A) binding protein cytoplasmic 1	(114)
PAX2	Paired box 2	(6)
PAX3	Paired Box 3	(36)
PAX8	Paired box 8	(114)
PBX4	PBX homeobox 4	(114)
PCSK2	Proprotein convertase subtilisin/kexin type 2	(114)
PDGFA	Platelet Derived Growth Factor Subunit A	(176)
PITX2	Paired Like Homeodomain 2	(107)
PLCXD1	Phosphatidylinositol specific phospholipase C X domain containing 1	(114)
PRKCE	Protein kinase C epsilon	(114)
PROP1	PROP paired-like homeobox 1	(177)
RARRES1	Retinoic acid receptor responder 1	(114)
RASD1	Ras related dexamethasone induced 1	(114)

Anexo 5: Genes candidatos avaliados neste estudo (conclusão)

Gene	Nome	Ref
RNASE6	Ribonuclease A family member k6	(114)
SFRP2	Secreted frizzled related protein 2	(114)
SHH	Sonic hedgehog	(6)
SIX1	SIX Homeobox 1	(178)
SLC26A4	Solute carrier family 26 member 4	(6)
SLC5A5	Solute carrier family 5 member 5	(6)
SMAD3	SMAD family member 3	(114)
SOX17	SRY-box 17	(6)
SOX2	SRY-Box Transcription Factor 2	(6)
SOX9	SRY-Box 9	(6)
TAL1	TAL BHLH Transcription Factor 1,	(7)
TAZ	Tafazzin	(6)
TBX1	T-box 1	(179)
TCAP	Titin-cap	(114)
TG	Thyroglobulin	(6)
TGFB2	Transforming Growth Factor Beta 2	(180)
THRA	Thyroid hormone receptor alpha	(159)
THRB	Thyroid hormone receptor beta	(159)
THRSP	Thyroid hormone responsive	(114)
TNFAIP2	TNF alpha induced protein 2	(114)
TNNC2	Troponin C2, fast skeletal type	(114)
TNXB	Tenascin XB	(181)
TPO	Thyroid peroxidase	(6)
TRA	T-cell receptor alpha locus	(114)
TRH	Thyrotropin releasing hormone	(6)
TSHb	Thyroid stimulating hormone beta	(6)
TSHR	Thyroid stimulating hormone receptor	(6)
TUBB1	Tubulin Beta 1 Class VI	(182)
TWSG1	Thioredoxin interacting protein	(183)
TXNIP	Thioredoxin interacting protein	(114)
TXNRD1	Thioredoxin Reductase 1	(184)
TYROBP	TYRO protein tyrosine kinase binding protein	(114)
VEGFA	Vascular endothelial growth factor A	(172)

Anexo 6: Genes regulados por FOXE1 (185) (continua)

Gene	Nome
Adamts9	ADAM Metallopeptidase With Thrombospondin Type 1 Motif 9
Ahcy	Adenosylhomocysteinase
Amigo3	Adhesion Molecule With Ig Like Domain 3
Ankrd37	Ankyrin repeat domain-containing protein 37
Atmin	ATM interactor
Bcam	Basal Cell Adhesion Molecule (Lutheran Blood Group)
Bet1	BET1 homolog
Casp4	Caspase-4
Cdh1	Cadherin 1
Coq10b	Coenzyme Q10B
Creld2	Cysteine Rich With EGF Like Domains 2
Crip2	Cysteine Rich Protein 2
Ctgf	Connective tissue growth factor
Ddit3	DNA Damage Inducible Transcript 3
Derl3	Derlin 3
Dnajb11	DnaJ Heat Shock Protein Family (Hsp40) Member B11
Dnajb9	DnaJ Heat Shock Protein Family (Hsp40) Member B9
Dnajc3	DnaJ Heat Shock Protein Family (Hsp40) Member C3
Duox2	Dual Oxidase 2
Dusp5	Dual Specificity Phosphatase 5
Dynlrb2	Dynein Light Chain Roadblock-Type 2
Elovl2	ELOVL Fatty Acid Elongase 2
Engase	Endo-Beta-N-Acetylglucosaminidase
Ero1lb	Endoplasmic Reticulum Oxidoreductase 1 Beta
Etv5	ETS Variant Transcription Factor 5
Fgf18	Fibroblast Growth Factor 18
Folr1	Folate Receptor Alpha
Ggct	Gamma-Glutamylcyclotransferase
Gmppb	GDP-Mannose Pyrophosphorylase B
HLA-DRA	Major Histocompatibility Complex, Class II, DR Alpha
Hsp90b1	Heat Shock Protein 90 Beta Family Member 1
Hspa5	Heat Shock Protein Family A (Hsp70) Member 5
Hyou1	Hypoxia Up-Regulated 1
Igf2bp2	Insulin Like Growth Factor 2 mRNA Binding Protein 2
Il23a	Interleukin 23 Subunit Alpha
Krt20	Keratin 20
LOC365444	
Manf	Mesencephalic Astrocyte Derived Neurotrophic Factor

Anexo 6: Genes regulados por FOXE1 (185) (conclusão)

Gene	Nome
Mfsd2	Major Facilitator Superfamily Domain Containing 2A
Nr4a2	Nuclear Receptor Subfamily 4 Group A Member 2
Nupr1	Nuclear Protein 1, Transcriptional Regulator
Pdia4	Protein Disulfide Isomerase Family A Member 4
Prima1	Proline Rich Membrane Anchor 1
Prss8	Serine Protease 8
Ril	Reversion-induced LIM
Riok3	RIO Kinase 3
S100a4	S100 Calcium Binding Protein A4
Sdf2l1	Stromal Cell Derived Factor 2 Like 1
Sec23b	SEC23 Homolog B, COPII Coat Complex Component
Sel1l	SEL1L Adaptor Subunit Of ERAD E3 Ubiquitin Ligase
Slit1	Slit Guidance Ligand 1
Tm4sf1	Transmembrane 4 L Six Family Member 1
Tmem140	Transmembrane Protein 140
Tmem66	Transmembrane Protein 66
Zfand2a	Zinc Finger AN1-Type Containing 2A

Anexo 7: Genes regulados por PAX8 (186, 187) (continua)

Gene	Nome
ACOT2	Acyl-CoA thioesterase 2
ACY1	Aminoacylase 1
ADAMTS9	ADAM metalloproteinase with thrombospondin type 1 motif 9
ALCAM	Activated leukocyte cell adhesion molecule
ANKRD9	Ankyrin repeat domain 9
ARHGAP22	Rho GTPase activating protein 22
BHLHE40	Basic helix-loop-helix family member e40
BRAF	B-Raf proto-oncogene, serine/threonine kinase
CAMK1G	Calcium/calmodulin dependent protein kinase IG
CAMKK2	Calcium/calmodulin dependent protein kinase 2
CD47	CD47 molecule
CDA	Cytidine deaminase
CDH16	Cadherin 16
CDH16	Cadherin 16
CFD	Complement factor D
CITED2	Cbp/p300 interacting transactivator with Glu/Asp rich carboxy-terminal domain 2
CRYAB	Crystallin alpha B
CTGF	Connective tissue growth factor
CXCL1	C-X-C motif chemokine ligand 1
DGAT2	Diacylglycerol O-acyltransferase 2
EGR1	Early growth response 1
EIF4E	Eukaryotic translation initiation factor 4E
ENPP1	Ectonucleotide pyrophosphatase/phosphodiesterase 1
F10	Coagulation factor X
FAM13A	Family with sequence similarity 13 member A
FGFR2	Fibroblast growth factor receptor 2
FOXE1	Forkhead box E1
GALK2	Galactokinase 2
GCSH	Glycine cleavage system protein H
GJA4	Gap junction protein alpha 4
GSTP1	Glutathione S-transferase pi 1
HACD4	3-hydroxyacyl-CoA dehydratase 4
HSD17B1	Hydroxysteroid 17-beta dehydrogenase 1
IGFBP5	Insulin like growth factor binding protein 5
IGFBP7	Insulin like growth factor binding protein 7
IRGQ	Immunity related GTPase Q
JUN	Jun proto-oncogene, AP-1 transcription factor subunit
KCNJ15	Potassium voltage-gated channel subfamily J member 15

Anexo 7: Genes regulados por PAX8 (186, 187) (conclusão)

Gene	Nome
KCNJ16	Potassium voltage-gated channel subfamily J member 16
KCNK1	Potassium two pore domain channel subfamily K member 1
KRT14	Keratin 14
KRT7	Keratin 7
LACTB	Lactamase beta
LRP8	LDL receptor related protein 8
LRRC58	Leucine rich repeat containing 58
NFKB1	Nuclear factor kappa B subunit 1
NR3C2	Nuclear receptor subfamily 3 group C member 2
NRIP3	Nuclear receptor interacting protein 3
NUP107	Nucleoporin 107
OPRK1	Opioid receptor kappa 1
OSTALPHA	Complement factor D
PBLD	Phenazine biosynthesis like protein domain containing
POMT1	protein O-mannosyltransferase 1
PRR5L	Proline rich 5 like
RAB17	RAB17, member RAS oncogene family
RASL10A	RAS like family 10 member A
RASSF2	Ras association domain family member 2
RSAD2	Radical S-adenosyl methionine domain containing 2
RUNX2	Runt related transcription factor 2
SLC26A7	Solute carrier family 26 member 7
SMIM22	Small integral membrane protein 22
SPARC	Secreted protein acidic and cysteine rich
STS	Steroid sulfatase
TAZ	Tafazzin
TEKT4	Tektin 4
TG	Thyroglobulin
TMEM140	Transmembrane protein 140
TRIB1	Tribbles pseudokinase 1
WBP2	WW domain binding protein 2
WNT4	Wnt family member 4

Research Article

Histopathological Characterization and Whole Exome Sequencing of Ectopic Thyroid: Fetal Architecture in a Functional Ectopic Gland from Adult Patient

Rosalinda Yasato Camargo,¹ Cristina Takami Kanamura,² Celso Ubirajara Friguglietti,³ Célia Regina Nogueira,⁴ Sonia Iorcansky,⁵ Alfio José Tincani,⁶ Ana Karina Bezerra,⁷ Ester Brust,^{8,9} Fernanda Christtanini Koyama,¹⁰ Anamaria Aranha Camargo,¹⁰ Fernanda Orpinelli R. Rego,¹⁰ Pedro Alexandre Favoretto Galante,¹⁰ Geraldo Medeiros-Neto,¹ and Ileana Gabriela Sanchez Rubio^{8,9}

¹Thyroid Unit, Cellular and Molecular Endocrine Laboratory, LIM-25, Faculdade de Medicina da Universidade de São Paulo (FMUSP), Avenida Doutor Arnaldo 455, Cerqueira César, 01246-904 São Paulo, SP, Brazil

²Adolfo Lutz Institute, São Paulo Public Health Service, Av. Dr. Arnaldo 355, Cerqueira César, 01246-000 São Paulo, SP, Brazil

³Head and Neck Surgery of Santa Catarina Hospital, Av. Paulista 200, Bela Vista, 01310-000 São Paulo, SP, Brazil

⁴Department of Internal Medicine, Botucatu School of Medicine, UNESP, Av. Prof. Montenegro, s/n Distrito de Rubião Junior, 18618-687 Botucatu, SP, Brazil

⁵Servicio de Endocrinología, Hospital de Pediatría Dr. Juan Garrahan, Combate de los Pozos 1881, C1245AAM Buenos Aires, Argentina

⁶Departamento de Cirurgia na Disciplina de Cirurgia de Cabeça e Pescoço da Faculdade de Ciências Médicas da UNICAMP, R. Tessália Vieira de Camargo 126, 13083-887 Campinas, SP, Brazil

⁷Medicine School, Universidade de Fortaleza (Unifor), Av. Washington Soares 1321, Edson Queiroz, 60811-905 Fortaleza, CE, Brazil

⁸Postgraduate Program in Biotechnology, Universidade Federal de São Paulo (UNIFESP), Pedro de Toledo 669, 040399-032 São Paulo, SP, Brazil

⁹Thyroid Molecular Sciences Laboratory, Universidade Federal de São Paulo, Departamento de Ciências Biológicas, Postgraduation Programs in Biotechnology and Structural and Functional Biology, UNIFESP, Pedro de Toledo 669, 040399-032 São Paulo, SP, Brazil

¹⁰Molecular Oncology Center, Hospital Sírio-Libanés, Rua Prof. Daher Cutait 69, 01308-060 São Paulo, SP, Brazil

Correspondence should be addressed to Ileana Gabriela Sanchez Rubio; ileana.rubio@unifesp.br

Received 30 August 2017; Accepted 16 November 2017; Published 8 February 2018

Academic Editor: Maria L. Dufau

Copyright © 2018 Rosalinda Yasato Camargo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ectopic thyroid results from a migration defect of the developing gland during embryogenesis causing congenital hypothyroidism. But it has also been detected in asymptomatic individuals. This study aimed to investigate the histopathological, functional, and genetic features of human ectopic thyroids. Six samples were histologically examined, and the expression of the specific thyroid proteins was assessed by immunohistochemistry. Two samples were submitted to whole exome sequencing. An oropharynx sample showed immature fetal architecture tissue with clusters or cords of oval thyrocytes and small follicles; one sample exhibited a normal thyroid pattern while four showed colloid goiter. All ectopic thyroids expressed the specific thyroid genes and T4 at similar locations to those observed in normal thyroid. No somatic mutations associated with ectopic thyroid were found. This is the first immature thyroid fetal tissue observed in an ectopic thyroid due to the arrest of structural differentiation early in the colloid stage of development that proved able to synthesize thyroid hormone but not to respond to TSH. Despite the ability of all ectopic thyroids to synthesize specific thyroid proteins and T4, at some point in life, it may be insufficient to support body growth leading to hypothyroidism, as observed in some of the patients.

REVIEW

Congenital hypothyroidism and thyroid cancer

Gustavo Penna^{1,2}, Ileana G S Rubio^{3,4,5}, Ester Saraiva Brust^{3,4}, Juliana Cazarin⁶, Fabio Hecht⁶, Nina Ramalho Alkmim^{2,7}, Kamilla M A Brandão Rajão^{2,8} and Helton Estrela Ramos^{9,10,11}

¹Department of Clinical Medicine, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

²Endocrinology Service of Instituto Orizonti, Belo Horizonte, Minas Gerais, Brazil

³Department of Biological Science, Molecular Thyroid Science Laboratory, Federal University of São Paulo, São Paulo, Brazil

⁴Biotechnology Post-graduation Program, Federal University of São Paulo, São Paulo, Brazil

⁵Estructural and Functional Biology Program, Federal University of São Paulo, São Paulo, Brazil

⁶Laboratory of Endocrine Physiology, Institute of Biophysics Carlos Chagas Filho, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

⁷Post-Graduation Program Molecular Medicine, Faculty of Medicine, Federal University of Minas Gerais, Belo Horizonte, Brazil

⁸Endocrinology Service of Hospital das Clínicas, Federal University of Minas Gerais, Belo Horizonte, Brazil

⁹Department of Bioregulation, Health & Science Institute, Thyroid Study Laboratory, Federal University of Bahia, Salvador, Brazil

¹⁰Post-Graduation Program Medicine and Health, Faculty of Medical Sciences, Federal University of Bahia, Salvador, Bahia, Brazil

¹¹Post-Graduate Program in Interactive Processes of Organs and Systems, Health & Science Institute, Federal University of Bahia, Salvador, Bahia, Brazil

Correspondence should be addressed to G Penna: gustavocpenna@gmail.com

Abstract

Differentiated thyroid carcinoma (DTC) combined with congenital hypothyroidism (CH) is a rare situation, and there is no well-established causal relationship. CH is a common congenital endocrine, while DTC occurring in childhood represents 0.4–3% of all malignancies at this stage of life. The association of CH with DTC could be related to dysmorphogenetic goiter (DHG) or developmental abnormalities. This review will explore the clinical features and the molecular mechanisms potentially associated with the appearance of DTC in CH: sporadic somatic driver mutations, chronic increase of thyroid-stimulating hormone (TSH) levels, higher concentrations of hydrogen peroxide (H₂O₂), cell division cycle associated 8 (*Borelain/CDC8*) gene mutations, and in others genes associated with CH – either alone or associated with the mechanisms involved in dysmorphogenesis. There are some pitfalls in the diagnosis of thyroid cancer in patients with CH with nodular goiter, as the proper cytological diagnosis of nodules of patients with dysmorphogenesis might be demanding due to the specific architectural and cytological appearance, which may lead to an erroneous interpretation of malignancy. The purpose of this article is to suggest an analytical framework that embraces the fundamental relationships between the various aspects of CH and CDT. In face of this scenario, the entire genetic and epigenetic context, the complex functioning, and cross talk of cell signaling may determine cellular mechanisms promoting both the maintenance of the differentiated state of the thyroid follicular cell and the disruption of its homeostasis leading to cancer. Whereas, the exact mechanisms for thyroid cancer development in CH remain to be elucidated.

Key Words

- ▶ congenital hypothyroidism
- ▶ goiter
- ▶ dysmorphogenesis
- ▶ thyroid cancer

Endocrine-Related Cancer
(2021) 28, R217–R230

Bibliografia consultada

Normas para teses e dissertações [Internet]. 3ª ed. São Paulo: Universidade Federal de São Paulo, Biblioteca Antônio Rubino de Azevedo, Coordenação de Cursos; 2021 [cited 2021 Feb 10]. Available from: <https://www.bibliotecacsp.unifesp.br/servicos/normas-teses-dissetecoes>.