The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Collection And Analysis of PAX8 Variants and Their Impact on Protein Function Via ClinVar and AlphaMissence

Twisha Sharma

Valley Christian High School

Dr Vera Khodzhaeva

Acknowledgements

This research article has been written under the guidance and mentorship of Dr. Vera Khodzhaeva, research associate at the University of Cambridge, UK. Cambridge Centre for International Research assisted me in the process of publication.

Abstract

This study analyzes the genetic influences of paired box gene 8 (PAX8) on Hashimoto's thyroiditis (HT) by examining PAX8 protein structure and function. Using a computational approach by comparing data from ClinVar and AlphaMissence databases, PAX8 variants and their impact on protein function were classified. The research identified 33 PAX8 genetic variants associated with HT, focusing on their pathogenicity. ClinVar was used to document variants, while AlphaMissense provided an additional layer of classification. Molecular models were created to visualize how these mutations affect protein structure, using data from Protein Data Bank and AlphaFold, refined through PyMol and InkScape. Findings revealed that 62.5%

of pathogenic mutations were concentrated in the PAI subdomain of the Paired Box Domain, with a smaller portion in the RED subdomain. These regions are crucial for DNA recognition, and mutations here likely disrupt the gene's ability to activate promoters essential for thyroid function. Interestingly, ClinVar and AlphaMissense sometimes conflicted in their classifications but agreed on many pathogenic variants. This study highlights the critical role of PAX8 mutations in HT development, particularly in protein stabilization. By improving the understanding of these genetic influences, this research could contribute to more precise diagnostic tools and targeted treatments for HT.

Keywords: Hashimoto's thyroiditis, PAX8, gene variants, ClinVar, AlphaMissence

Introduction

Hashimoto's thyroiditis (HT), or hypothyroidism, is a prevailing autoimmune thyroid disease first described by Dr. Hakaru Hashimoto in 1912 [1]. The condition most prominently affects women, causing hormone imbalance and thyroid dysfunction, and shows increased incidence with age [2, 3]. HT is caused by the production of antithyroid antibodies that destroy tissues in the thyroid, leading to the symptoms of the condition [4]. Symptoms range from asymptomatic to severe hypothyroidism in those affected and can include constipation, difficulty concentrating or thinking, dry skin, fatigue, hair loss, depression, cold intolerance, heavy or irregular menstrual period, and an enlarged neck or the presence of a goiter [1, 2].

Diagnosis can be tricky, and currently involves clinical data such as medical and family history and a physical exam along with blood tests to check for high levels of thyroid stimulating hormone (TSH) and low levels of thyroxine [4]. Standard treatment for HT involves levothyroxine replacement therapy at about 1.4 to 1.8 mcg/kg/day, which helps to convert the thyroid hormone into its active form [2, 4]. Triiodothyronine (T3) therapy is emerging as a

potential treatment for HT, but there is still currently not enough evidence to support or oppose it [2, 4].

People diagnosed with HT are more likely to develop cardiovascular issues such as high blood pressure, high cholesterol, heart disease and heart failure [2]. HT is also linked to myxedema, a condition where one's body systems slow until they become potentially fatal [5]. Problems during pregnancy and tumors may also develop. It is more likely for an individual to develop HT if they have other autoimmune disorders including type 1 diabetes, celiac disease, rheumatoid arthritis, and systemic lupus erythematosus [6]. Although about 70% of the risk of developing Hashimoto's thyroiditis is a result of one's genetic background, environmental factors that may influence an individual's chances include radiation exposure and an increased iodine intake [7].

Hashimoto's thyroiditis is affected by many genes, including HLA (human leukocyte antigen), CTLA-4 (Cytotoxic T Lymphocyte Antigen-4), PTPN22 (Protein Tyrosine Phosphatase Nonreceptor-Type 22), Tg (Thyroglobulin), Vitamin D receptors, and cytokine genes [8]. These genes perform a variety of functions. HLA provides the code for producing human leukocyte antigens, while CTLA-4 and PTPN22 function in regulating the immune system [8]. Tg codes for an antigen specifically manufactured in the thyroid that spreads through the bloodstream [8]. Cytokine genes and vitamin D receptors are also utilized in regulating the immune response [8].

This paper aims to discuss the genetic influence on HT with a focus on the PAX8 gene. The PAX8 gene produces a protein whose activity is regulated in the N-terminal, which has three domains, with a binding site in the Paired Box Domain [9]. The gene is of the PAX8 superfamily, providing a transcription factor [9]. The PAX8 gene is utilized in both embryos and adults, the former in helping differentiate organs and the latter in triggering thyroid hormone

synthesis [9]. The gene itself is regulated with two types of transcription factors: thyroid transcription factor-1 and thyroid transcription factor-2 and Hex (a transcription repressor) [9, 10].

Isoforms of the protein produced from the PAX8 gene regulate thyroglobulin, thyrotropin receptor genes, sodium/iodide symporter, and thyroperoxidase production [9]. Many variants found in the PAX-8 gene (specifically in the Paired Box Domain) have been previously connected to thyroid dysgenesis, a condition that leads to congenital hypothyroidism [9]. Congenital hypothyroidism is a condition where thyroid hormone production is deficient from birth, and can be goitrous or nongoitrous (without a goiter) [11].

Single amino acid substitutions might affect protein folding and function [27]. Being able to differentiate between benign and pathogenic variants is essential for treatment strategies [27]. Collecting information on the mutations involved in HT condition can lead to improved treatment and understanding of the disease mechanism [27]

Methods

Several databases were utilized to produce results and accurate analysis. ClinVar, a public archive perpetuated by the National Institutes of Health (NIH), is a freely available website displaying genetic variants of humans and the impact of these mutations on conditions [3].

ClinVar was employed to investigate variants associated with Hashimoto's Thyroiditis, Hypothyroidism, and Congenital Hypothyroidism. Nongoitrous congenital hypothyroidism was included and all found mutations were documented in a table (Appendix Table 1). These variants were then organized, examined, and linked to corresponding publications through PubMed, a

free, public resource maintained by the National Library of Medicine (NLM); the website provides a reputable source for biomedical literature [12].

It was then perceived that several genes were more closely linked to HT as compared to others. From the list of impactful genes, PAX8 had the largest amount of genetic mutations associated with HT, therefore it was selected for further analysis. The variants were also linked to publications if applicable and their locations and effects on protein function were noted. Protein coding variants were then placed into another database known as AlphaMissence. AlphaMissence is a pathogenicity predicting website for amino acid substitutions developed by Jun Cheng and colleagues [13]. Intron/UTR and exon variants were collected together into a separate table (Supplementary Table 1). The exon variants were separated, identified and analyzed for pathogenicity predictions, and the data collected was edited, arranged, and added to the existing table (Table 1).

The PAX8 gene was located in the Protein Data Bank (PDB), a free global website containing thousands of structures of biological molecules, and downloaded into PyMol [14]. PyMol is a protein graphic program that allows for one to present, manipulate, and create high quality images of molecular protein structures [15].

Each protein-coding variant was consequently highlighted in colors corresponding to its pathogenicity. Alpha helices were coded in a pale cyan tint. Blue indicated variants that were classified as uncertain or conflicting pathogenicity between AlphaMissense ClinVar. TV orange correlated to variants that were uncertain or conflicting in ClinVar, but showed as likely pathogenic in AlphaMissence. Bright red connected to variants that were pathogenic in both databases. Dark gray referred to mutations that were uncertain or conflicting classifications in ClinVar, but likely benign classifications in AlphaMissence. Light red or pale pink indicated

likely pathogenic variants in both databases. Purple specified mutations that were likely benign in both databases. Lastly, sand (a dull shade of yellow) indicated variants that were likely benign in ClinVar and likely pathogenic in AlphaMissence. The terminals Cys45 and Cys57 were colored in red text and colored magenta (to represent the regulator) and dark teal (to represent a regulator that also has a corresponding pathogenic mutation).

Several of the colored variants were found to be located in regions past that displayed in the PDB structure available. To solve the issue, a website known as AlphaFold was used. AlphaFold is a computational algorithm that utilizes machine learning to create protein structure predictions of biological molecules [16]. A relatively accurate model of the PAX8 paired box protein was selected and placed into PyMol, where the remaining mutations were identified and colored accordingly.

Both protein models from PyMol were then exported as portable network graphics (PNGs) and imported into a software known as Inkscape, a free computer program that allows for professional rendering and high quality image exportation [17]. All of the mutations from both of the models (from PDB and AlphaFold) were labeled and a key depicting the representation of each color was created, leading to the formation of the images within this paper (Figure 1-2).

Results

Throughout the data gathering process, it was crucial to consider and collect information about different genes and their variants that affect Hashimoto's thyroiditis. ClinVar database was chosen for data collection since it contains clinical data and variants for diseases and genes [3]. 12 genes affecting HT were identified (Supplementary Table 1). From the respective large number of genes, the gene with the largest amount of mutations was selected for further examination to predict how the variants affected its function. This gene was found to be PAX8, and all subsequent variants were examined and compiled into a table (Table 1). Next, only exon variants were analyzed further. 33 PAX8 exon variants were analyzed (Table 1). An existing computational method, AlphaMissence, was utilized in an effort to predict functional significance of these variants [13]. AlphaMissence is a database that uses artificial intelligence in order to classify variant pathogenicity [13]. Comparison between Clinvar and AlphaMissence was documented in Table 1.

Table 1

PAX8 Exon Variants associated with Hashimoto's Thyroiditis (hypothyroidism, congenital, nongoitrous) According to ClinVar and AlphaMissence Pathogenicity Classifications

DNA Base	Amino Acid Substitution	Type	ClinVar	AlphaMissence	Publications
c.985T	p.Phe329Leu	Single nucleotide variant (missense variant)	Benign/Likely benign	Likely Pathogenic	Carvalho A et al, 2013

c.1139C	p.Thr380Ile	Single nucleotide variant (missense variant)	Likely benign	likely benign	No publications found in the database
c.215G	p.Arg72Gln	Single nucleotide variant (missense variant)	Likely benign	likely pathogenic	No publications found in the database
c.1275G	p.Leu425 (shows many variants in AlphaMissence)	Single nucleotide variant (synonymous variant)	Uncertain significance	varying significance	No publications found in the database
c.1267A	p.Ser423Gly	Single nucleotide variant (missense variant)	Uncertain significance	likely benign	No publications found in the database
c.1226C	p.Pro409Arg	Single nucleotide variant (missense variant)	Uncertain significance	uncertain significance	No publications found in the database
c.1189G	p.Gly397Arg	Single nucleotide variant (missense variant)	Uncertain significance	likely pathogenic	No publications found in the database
c.1185G	p.Val395 (shows many variants in AlphaMissence)	Single nucleotide variant (synonymous variant)	Uncertain significance	varying significance	No publications found in the database

c.1060C	p.Gln354Glu	Single nucleotide variant (missense variant)	Uncertain significance	likely benign	No publications found in the database
c.297C	p.Asn99 (shows many variants in AlphaMissence)	Single nucleotide variant (synonymous variant)	Uncertain significance	likely pathogenic	No publications found in the database
c.858G	p.Gly286 (shows many variants in AlphaMissence)	Single nucleotide variant (synonymous variant)	Uncertain significance	likely benign or ambiguous	No publications found in the database
c.817G	p.Asp273Asn	Single nucleotide variant (missense variant)	Uncertain significance	likely benign	No publications found in the database
c.324A	p.Arg108 (shows many variants in AlphaMissence)	Single nucleotide variant (synonymous variant)	Uncertain significance	likely pathogenic	No publications found in the database
c.280G	p.Asp94Asn	Single nucleotide variant (missense variant)	Uncertain significance	likely benign	No publications found in the database
c.441C	p.Cys147 (shows many variants in AlphaMissence)	Single nucleotide variant (synonymous variant)	Uncertain significance	varying significance	No publications found in the database
c.1028A	p.Asn343Ser	Single nucleotide variant (missense variant)	Conflicting classifications of pathogenicity	likely benign	No publications found in the database

c.1009T	p.Ser337Ala	Single nucleotide variant (missense variant)	Conflicting classifications of pathogenicity	likely benign	Vizcaino MA et al, 2019
c.346G	p.Val116Ile	Single nucleotide variant (missense variant)	Conflicting classifications of pathogenicity	likely benign	No publications found in the database
c.237dup	p.Lys80fs (shows many variants in AlphaMissence)	Duplication (frameshift variant)	Likely pathogenic	Likely pathogenic	No publications found in the database
c.236C	p.Ser79Phe	Single nucleotide variant (missense variant)	Likely pathogenic	Likely pathogenic	No publications found in the database
c.205G	p.Gly69Ser	Single nucleotide variant (missense variant)	Likely pathogenic	Likely pathogenic	No publications found in the database
c.203C	p.Thr68Ile	Single nucleotide variant (missense variant)	Likely pathogenic	Likely pathogenic	No publications found in the database
c.160A	p.Ser54Cys	Single nucleotide variant (missense variant)	Likely pathogenic	Likely pathogenic	No publications found in the database
c.397C	p.Arg133Trp	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Vincenzi M et al, 2014

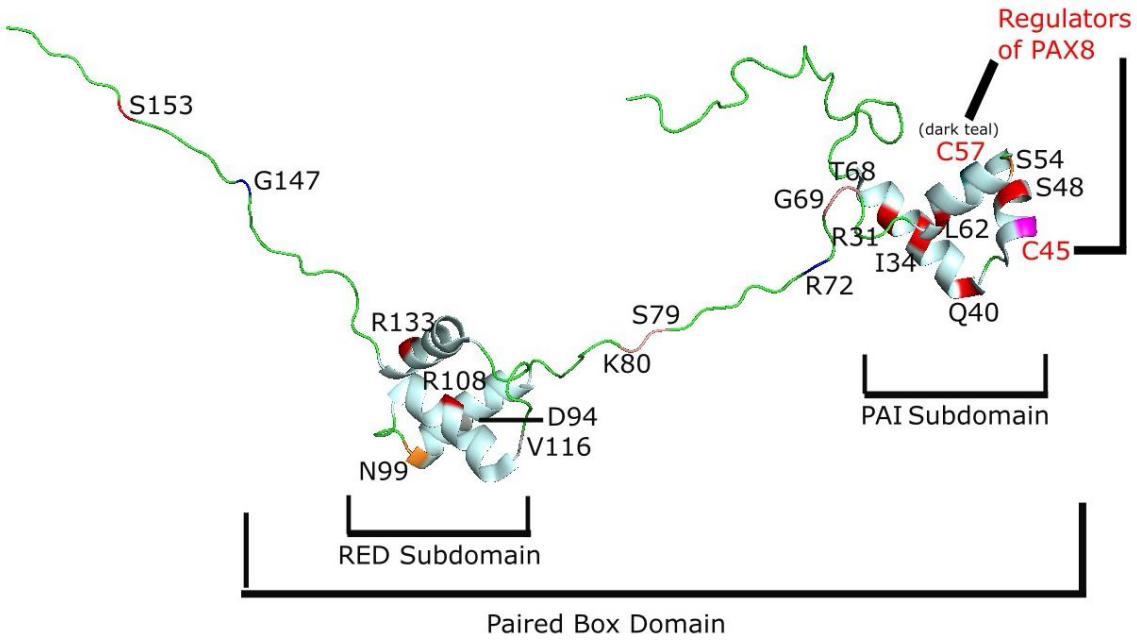
c.101T	p.Ile34Asn	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Lanzerath K et al, 2006
c.119A	p.Gln40Pro	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Congdon T et al, 2001
c.457_458del	p.Leu153 (shows many variants in AlphaMissence)	Deletion (frameshift variant)	Pathogenic	varying significance	No publications found in the database
c.322C	p.Arg108Ter	Single nucleotide variant (nonsense)	Pathogenic	no classification found in database	Macchia PE et al, 1998
c.185T	p.Leu62Arg	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Macchia PE et al, 1998
c.170G	p.Cys57Tyr	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Vilain C et al, 2001
c.160A	p.Ser54Gly	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Meeus L et al, 2004
c.143C	p.Ser48Phe	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Grasberger H et al, 2005

c.92G	p.Arg31His	Single nucleotide variant (missense variant)	Pathogenic	likely pathogenic	Macchia PE et al, 1998

A protein model of the Paired Box Domain (PABD) of PAX8, found using the protein data bank database (PDB), was uploaded into PyMol for manipulation in an attempt to examine the location of the variants (Figure 1). The PABD contains two subdomains; RED is the C-terminal of the PABD while PAI is the N-terminal [19]. Both subdomains are required for proper PAX8 protein recognition [19]. Each mutation and the alpha helices of the PABD were subsequently colored according to a color key (Figure 1). The regulating terminals of the PABD (Cys45 and Cys57) and all of the variants were also labeled.

Figure 1

PAX8 Paired Box Domain With Pathogenicity Color Coded Mutations

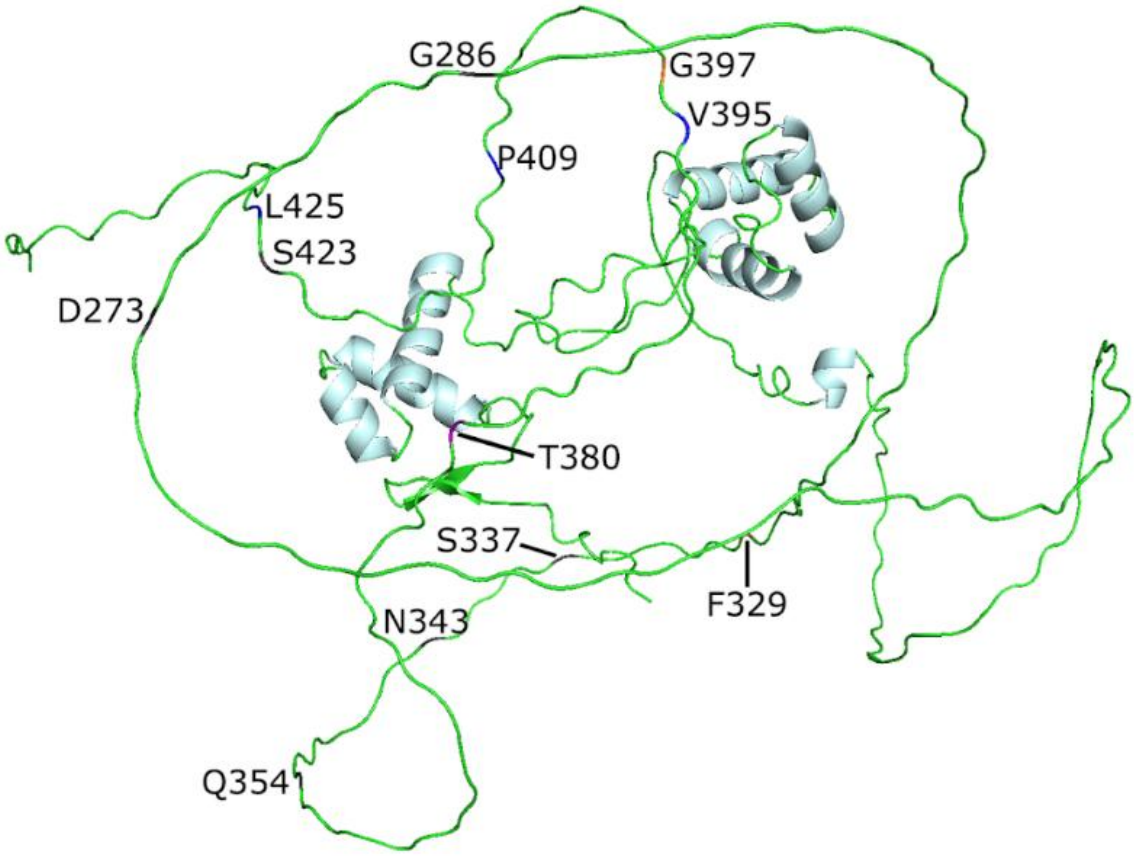


The solution structure of PAX8 was downloaded from PDB (2K27). Of subdomains shown above, RED is the C-terminal of the PABD while PAI is the N-terminal. All colors in the figure above refer to the pathogenicity of the variant they correspond to. Darker blue corresponds to variants which have uncertain or conflicting pathogenicities between AlphaMissence and ClinVar. Bright red refers to pathogenic variants from both databases, while light red correlates to likely pathogenic variants from both databases. Purple defines variants that are likely benign in both databases, while TV orange outlines variants that are uncertain or conflicting in ClinVar, but likely pathogenic in AlphaMissence. Similarly, dark gray illustrates variants that are uncertain or have conflicting classifications in ClinVar, but are shown to be likely benign in AlphaMissence. Alpha helices are shaded in pale blue, while a sand (dull yellow) color corresponds to mutations that are likely benign in ClinVar but likely pathogenic in AlphaMissence. Magenta (Cys45) represents a regulator while dark teal represents a regulator

terminal that is also the location of a pathogenic variant. The regulator terminals (Cys45 and Cys57) are marked with red text.

Figure 2

PAX8 Protein Structure Displaying Pathogenicity Color Coded Mutations Located >158 Amino Acids



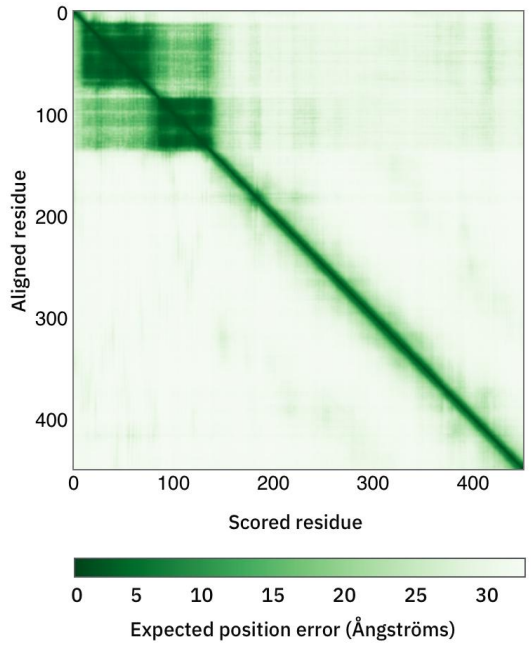
Protein structure was downloaded from AlphaFold (Q06710). All colors in the figure above refer to the pathogenicity of the variant they correspond to. Darker blue corresponds to variants which have uncertain or conflicting pathogenicities between AlphaMissence and ClinVar. Bright red refers to pathogenic variants from both databases, while light red correlates to likely pathogenic variants from both databases. Purple defines variants that are likely benign

in both databases, while TV orange outlines variants that are uncertain or conflicting in ClinVar, but likely pathogenic in AlphaMissence. Similarly, dark gray illustrates variants that are uncertain or have conflicting classifications in ClinVar, but are shown to be likely benign in AlphaMissence. Alpha helices are shaded in pale blue, while a sand (dull yellow) color corresponds to mutations that are likely benign in ClinVar but likely pathogenic in AlphaMissence. Magenta (Cys45) represents a regulator while dark teal represents a regulator terminal that is also the location of a pathogenic variant.

The PDB file of PAX8 did not have the full structure of the protein. Therefore, a model from AlphaFold was uploaded and mutations were labeled accordingly (Figure 2). AlphaFold utilizes artificial intelligence (AI) to conjecture the protein shape [15]. This process does not guarantee complete accuracy in the AlphaFold projection shown, so the per-residue model confidence score (pLDDT) demonstrates how confident the software is in its prediction by residue (Figure 3).

Figure 3

AlphaMissence pLDDT Chart Corresponding to Figure 2



Per-residue model confidence score (pLDDT) was downloaded from AlphaFold for PAX8 (Q06710).

Gene variants can be classified into several groups. Varying significance or conflicting classification refers to a variant whose pathogenicity is controversial. Uncertain significance refers to a variant whose pathogenicity is unknown; there may not be enough scientific research to confirm or deny the link between diseases, or studies may be conflicting. Benign or likely benign refers to a variant which is not disease-causing or likely not disease-causing. Pathogenic or likely pathogenic refers to variants that are disease-causing.

33 variants for the PAX8 protein were found (Table 1). Before examining their classifications, it is important to note that the ClinVar database is heavily based on clinical data, while the AlphaMissence database utilizes AI programs to predict pathogenicity [3, 13]. Variant Phe329Leu was observed to be benign or likely benign according to ClinVar and likely pathogenic according to AlphaMissence. Three variants Val116Ile, Asn343Ser, and Ser337Ala

had conflicting classifications of pathogenicity in ClinVar, but were marked as likely benign in AlphaMissence, while five others Ser423Gly, Gln354Glu, Asp273Asn, Asp94Asn, and Gly286 were uncertain in ClinVar and likely benign in AlphaMissence, with Gly286 being additionally ambiguous in AlphaMissence. Thr380Ile was classified to be likely benign in both databases. Three variants Leu425, Val395, and Cys147 were of uncertain significance in ClinVar, but had varying significance in AlphaMissence. Pro409Arg was of uncertain significance in both databases. Gly397Arg, Asn99, and Arg108 were uncertain in ClinVar, but predicted to be likely pathogenic in AlphaMissence. Arg72Gln was controversial, being likely benign in ClinVar and likely pathogenic in AlphaMissence. Five variants Lys80fs, Ser79Phe, Gly69Ser, Thr68Ile, and Ser54Cys were likely pathogenic in both databases. Leu153 was pathogenic in ClinVar and had varying significance in AlphaMissence, while Arg108Ter was marked to be pathogenic in ClinVar and yet was not documented in AlphaMissence. Arg133Trp, Ile34Asn, Gln40Pro, Leu62Arg, Cys57Tyr, Ser54Gly, Ser48Phe, and Arg31His were marked as pathogenic in ClinVar and likely pathogenic in AlphaMissence.

Following the locating of variants on protein structure (Figure 1-2), it was observed that the majority of the pathogenic mutations found were located in the alpha helices of the PAI subdomain, with two being found in the RED subdomain and one being found at Leu153. 62.5% (# of pathogenic mutations located inside of the PAI subdomain / # of total pathogenic mutations x 100) of the red colored pathogenic variants were located in the PAI subdomain. Variants with uncertain or conflicting classifications were mostly located outside of the known structure of the PAX8 protein, with the exception of variants R72 and G147. Likely benign variants in at least one database show a similar trend, with only D94 located within the PABD.

All variants categorized as uncertain or with conflicting pathogenicity were additionally evaluated for their pathogenicity based on their location in the PAX8 protein structure (Table 2). As variants located in alpha helices were largely noted to be pathogenic in both databases, uncertain or conflicting variants located in such alpha helices were additionally predicted to be likely pathogenic. Variants located in intrinsically displaced (IDP) regions were evaluated to be likely benign due to a similar pattern of benign variants being situated in these areas.

Table 2

PAX8 Variants Classification Based on Protein Structure Localization Analysis

DNA Base	Amino Acid Substitution	ClinVar	AlphaMissence	Protein Structure	Predicted Pathogenicity
c.1028A	p.Asn343Ser	Conflicting classifications of pathogenicity	likely benign	IDP region outside of known PAX8 structure	Likely benign
c.1009T	p.Ser337Ala	Conflicting classifications of pathogenicity	likely benign	IDP region outside of known PAX8 structure	Likely benign
c.1275G	p.Leu425 (shows many variants in AlphaMissence)	Uncertain significance	varying significance	IDP region outside of known PAX8 structure	Likely benign
c.1267A	p.Ser423Gly	Uncertain significance	likely benign	IDP region outside of known PAX8 structure	Likely benign

c.1226C	p.Pro409Arg	Uncertain significance	uncertain significance	IDP region outside of known PAX8 structure	Likely benign
c.1189G	p.Gly397Arg	Uncertain significance	likely pathogenic	IDP region outside of known PAX8 structure	Likely benign
c.1185G	p.Val395 (shows many variants in AlphaMissence)	Uncertain significance	varying significance	IDP region outside of known PAX8 structure	Likely benign
c.1060C	p.Gln354Glu	Uncertain significance	likely benign	IDP region outside of known PAX8 structure	Likely benign
c.346G	p.Val116Ile	Conflicting classifications of pathogenicity	likely benign	IDP region within the PABD.	Likely benign
c.297C	p.Asn99 (shows many variants in AlphaMissence)	Uncertain significance	likely pathogenic	Alpha helix in RED subdomain	Likely pathogenic
c.858G	p.Gly286 (shows many variants in AlphaMissence)	Uncertain significance	likely benign or ambiguous	IDP region outside of known PAX8 structure	Likely benign

c.817G	p.Asp273Asn	Uncertain significance	likely benign	IDP region outside of known PAX8 structure	Likely benign
c.324A	p.Arg108 (shows many variants in AlphaMissence)	Uncertain significance	likely pathogenic	Alpha helix in RED subdomain	Likely pathogenic
c.280G	p.Asp94Asn	Uncertain significance	likely benign	Alpha helix in RED subdomain	Likely pathogenic
c.441C	p.Cys147 (shows many variants in AlphaMissence)	Uncertain significance	varying significance	IDP region within the PABD.	Likely benign

Discussion

Hashimoto's thyroiditis (HT) is a prominent autoimmune thyroid disease that causes hormone imbalance and thyroid dysfunction [1, 2, 3]. A search of the ClinVar database showed that the following genes CEP128, CTLA4, DUOX2, IGSF1, IRS4, LOC126806316, PAX8, NKX2-5, TG, TPO, TSHR, and ZFAT are involved in Hashimoto's Thyroiditis (Supplementary Table 1). There are several variants identified for each gene. These variants are changes in the genomic sequence that may or may not influence the activity of the resulting protein [27]. Knowledge about variants can help to understand the foundation of a disease and alleviate the pain of those suffering from it by targeting the specific error in the body [27]. The focus of this research is on the PAX8 gene, as it displayed the largest amount of variants of the genes analyzed. The 33 variants examined within the bounds of this paper are all presented in the PAX8 protein [27].

It is known that the PAI and RED subdomains in the PAX8 proteins are its N and C terminals respectively, and interact together to allow a proper DNA recognition [19]. It was found that the PAI subdomain approximately binds to bases 7-15 of the binding site of the Tg gene promoter, while the RED subdomain approximately binds to bases 16-24. The RED subdomain in particular is essential for proper recognition of the PAI subdomain and it interacts first with the DNA during the process of binding [19].

Figure 1 shows all found exon mutations of PAX8 labeled and colored for pathogenicity; it was perceived that the majority of pathogenic mutations are located in the alpha helices of these subdomains. Other proteins have also similarly been influenced by mutations in their alpha helices. For example, ion transporters have reported to show different pathogenic variants within their alpha helices, and five different protein classes have shown pathogenic variants

congregated in their alpha helices as well [30]. Additionally, p.Arg156Pro, a variant located in an alpha-helix of proteins lamin A and lamin C, has been suggested as a cause of hereditary dilated cardiomyopathy [31]. It is also known that alpha helices tend to carry more mutations than beta strands [32].

It is likely that the mutations interfere with DNA recognition of the PAX8 protein. PAX8 protein recognition to DNA is vital to ensure proper regulation of thyroperoxidase, thyroglobulin, thyrotropin receptor genes, and sodium/iodide symporter production [9]. One mutation Cys57Tyr is located on C57, a residue that constitutes one of the two regulatory terminals in the PAI subdomain [9]. This mutation is pathogenic, pointing to signs that the condition is related to DNA recognition of the PAX8 protein; a study previously suggested that Cys57Tyr interferes with binding, causing TPO promoter affiliated gene transcription to become impractical [21].

It is known that Gln40Pro, the pathogenic variant, impairs the protein's ability to bind to a response element and hinders trans-activation of a thyroid peroxidase promoter luciferase reporter gene [20]. The variant is located in exon 3 and displays a heterozygous transversion of 119A to C, leading to the conserved glutamine at the carboxyl terminus of the first alpha helix of the paired box domain to be replaced with proline [20]. Cys57Tyr is yet another pathogenic variant that has been speculated to severely interfere with the DNA recognition of the PAX8 protein and thus make impossible the transcription of a TPO promoter affiliated gene [21]. The mutation is found within the PAI subdomain at the verge of the third alpha helix [21]. The third alpha helix is particularly used to recognize DNA, and directly connects to the DNA binding site at base pairs 4-8. Cys57 itself fits with a thymidine and guanine through Van der Waals interactions and hydrogen bonds respectively at base pairs 7 and 8 [21]. Pathogenic variant

Ser54Gly displays an inability for the mutant protein to bind to its thyroperoxidase gene promoter [22]. Such a variant also suffers the loss to transactivate thyroglobulin promoters or enhancers due to an incapability to work with Titf1 [22]. The glycine replaces a highly conserved serine found in all investigated PAX proteins, located in the amino-terminal homeodomain motif linking the second and third alpha helices of the PAI subdomain [22]. Arg108Ter is indicated to be a pathogenic *de novo* mutation found in exon 3 that truncates the paired box domain early in the RED subdomain, leading to high levels of Tg [23]. Leu62Arg, another pathogenic variant, supports a strong connection to thyroid gland reduction [23]. Pathogenic variant Arg31His, located in exon 2, resulted in congenital hypothyroidism (thyroid hypoplasia), from very high levels of TSH and very low levels of T4 [23]. Phe329Leu also displays an inability to bind to the TPO promoter binding DNA sequence; however is classified as benign or likely benign in ClinVar and likely pathogenic in AlphaMissence [24]. The variant is located in exon 2 [24]. It has also been suggested in the past that exon 3 of the PAX8 gene plays a significant role in hypothyroidism, where proper transcription and translation of the protein are hindered from variants [25].

It is known that Thr380Ile is classified as likely benign in both ClinVar and AlphaMissence, and is located in the intrinsically disordered protein (IDP) region beyond the known structure of the PAX8 protein (Figure 2). Mutations classified as uncertain in ClinVar but described as likely benign in AlphaMissence are D94, Q354, N343, S337, D273, S423, G286, and F329. The mutations are all located outside of the PABD except for D94, which is located within an alpha helix of the RED subdomain (Figure 2). Variants with uncertain or conflicting pathogenicity in ClinVar and AlphaMissence include G147, R72, L425, P409, and V395. Of this, only R72 is located within the PABD (Paired Box Domain). Variants depicted as uncertain

or conflicting in ClinVar but likely pathogenic in AlphaMissence include N99, S54, and G397. Of this, G397 is located outside of the PABD while N99 and S54 reside with the PABD. Mutations that were found to be likely pathogenic in both databases include N99, S54 (both are within the PABD) and G397. The location of such variants, the majority of the uncertain, varying, and likely benign mutations being outside of the PABD implies that these regions are less involved in the DNA binding of the PAX8 protein and mutations located here are therefore less likely to significantly affect protein function and structure. However, the existence of likely pathogenic mutation G397 being located outside the PABD demonstrates how these regions do affect the overall protein function despite not being affiliated with the PABD themselves.

The PAX8 protein itself is thought to be a transcription factor related to the activation of Tg and TPO genes that binds to their promoter sites and is compulsory for the survival of thyroid precursor cells [26, 28]. As mentioned above, the PAI and RED subdomains themselves cooperate together to correctly bind to DNA [19]. One can observe how these mutations all interfere with the PAX8 protein's ability to recognize DNA. Mutations located in the alpha helices especially are more likely to be pathogenic than those located in the IDP region of the PAX8 protein. This connecting fact was also the basis for the classification of variants in Table 2, where variants located in the IDP regions were listed as likely benign and those in the alpha helices as likely pathogenic.

These results are consistent with these previous studies that have shown how symptoms of Hashimoto's Thyroiditis or hypothyroidism have been linked to PAX8's inability to correctly bind to DNA [20, 21, 22, 23, 24, 25]. These findings could provide connecting insights between existing theories about HT and its connection to the PAX8 protein. The project was intended to coherently correlate and compile findings between PAX8 and HT and form a defined connection

with reason as to how and why the protein affects the disease. It is important to acknowledge that the disease is at least somewhat partially related to the mutant type PAX8's inability to bind correctly to DNA and stimulate TPO and Tg production. Thus, treatments of HT assisting PAX8 in this binding would likely be more effective. In the future, experiments on the variants documented in this paper and their effect on the protein's function could be considered.

This work also demonstrates the differences between ClinVar and AlphaMissence, two publicly used databases. ClinVar is dependent on clinical data while AlphaMissence predicts variant pathogenicity through the use of artificial intelligence [3, 13]. ClinVar provides more information per variant including conditions, genes, while AlphaMissence predicts pathogenicity of missense variants [3, 13]. Pathogenic variant Arg108Ter in ClinVar was not found in the AlphaMissence database.

It shall be noted that this paper has its limitations as there were no means to experimentally check the variants. All data was taken from the databases ClinVar and AlphaMissence. This means conclusions rely on reported clinical data and computational predictions. Therefore, conclusions drawn have not been validated with experimental data for any variant in this paper. The original PDB structure that has been experimentally determined displayed largely only the Paired Box Domain of the protein. Due to some variants being located further away in the protein than available in the PDB protein structure, AlphaFold was utilized for a predictional structure of the PAX8 gene, a website that employs artificial intelligence to estimate a protein's structure [14]. This allowed for the visualization of the variants located further away. Although AlphaFold utilizes advanced machine-learning algorithms, its predictions are not guaranteed to match the actual protein's structure. The AlphaFold structure was utilized within the bounds of this paper to visualize all of the variants, but no conclusion can be drawn

related to protein structure from the AlphaFold generated models that is reliable enough to use in future papers. However, the Paired Box Domain PDB structure is far more accurate and can be utilized in future papers.

Additionally, this paper highlights a highly probable interpretation of the collected and analyzed data, but requires test data to confirm listed conclusions. Future research should incorporate biochemical or cellular methods of determining each variant's inability to bind to the PAX8 domain and the effect of mutated PAX8 on Hashimoto's Thyroiditis. Such research may confirm the hypotheses presented in this paper and aid in determining the role of the PAX8 gene in the disease of Hashimoto's Thyroiditis. Examining how thyroid hormone synthesis or regulation is affected by the mutated PAX8 gene is also a promising method of understanding PAX8's relevance to Hashimoto's Thyroiditis. It is also advised to utilize and experimentally check the methods in this paper to other genes that play roles in HT. By understanding each gene's effect on HT, it is easier to find and target the exact issue when determining treatment strategies or creating drugs to alleviate the symptoms of or cure HT entirely.

Conclusion

This study shows the significant role of PAX8 gene and its mutations in Hashimoto's thyroiditis. By comparing data from ClinVar and AlphaMissence, the location of various PAX8 mutations and their effect on protein structure and function were better understood. Comparison between two databases ClinVar and AlphaMissence showed differences in pathogenicity classification. The findings reinforce the hypothesis that the clinical results of HT can be partially caused by the PAX8 protein being unable to properly bind to DNA. This research helps to connect what has already been learned regarding PAX8 and HT, lay a foundation for future investigations into the results of these mutations and provide a start for more targeted treatment and diagnosis.

References

Caturegli, P., de Remigis, A., Chuang, K., Dembele, M., Iwama, A., & Iwama, S. (2013). Hashimoto's thyroiditis: celebrating the centennial through the lens of the Johns Hopkins hospital surgical pathology records. *Thyroid : Official Journal of the American Thyroid Association*, *23*(2), 142–150. https://doi.org/10.1089/thy.2012.0554

Klubo-Gwiezdzinska, J., & Wartofsky, L. (2022b). Hashimoto thyroiditis: an evidence-based guide to etiology, diagnosis and treatment. *Polish Archives of Internal Medicine*, *132*(3). https://doi.org/10.20452/pamw.16222

Landrum, M. J., Chitipiralla, S., Brown, G. R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., Lyoshin, V., Maddipatla, Z., Maiti, R., Mitchell, J., O'Leary, N., Riley, G. R., Shi, W., Zhou, G., Schneider, V., … Kattman, B. L. (2020). ClinVar: improvements to accessing data. *Nucleic Acids Research*, *48*(D1), D835–D844. https://doi.org/10.1093/nar/gkz972

Mincer, D. L., & Jialal, I. (2024). *Hashimoto Thyroiditis*.

Chaker, L., Bianco, A. C., Jonklaas, J., & Peeters, R. P. (2017). Hypothyroidism. *Lancet (London, England)*, *390*(10101), 1550–1562. https://doi.org/10.1016/S0140-6736(17)30703-1

Ruggeri, R. M., Trimarchi, F., Giuffrida, G., Certo, R., Cama, E., Campennì, A., Alibrandi, A., de Luca, F., & Wasniewska, M. (2017). Autoimmune comorbidities in Hashimoto's thyroiditis: different patterns of association in adulthood and childhood/adolescence. *European Journal of Endocrinology*, *176*(2), 133–141. https://doi.org/10.1530/EJE-16-0737

Brent, G. A. (2010). Environmental exposures and autoimmune thyroid disease. *Thyroid : Official Journal of the American Thyroid Association*, *20*(7), 755–761. https://doi.org/10.1089/thy.2010.1636

Zaletel, K., & Gaberšček, S. (2011). Hashimoto's Thyroiditis: From Genes to the Disease. *Current Genomics*, *12*(8), 576–588. https://doi.org/10.2174/138920211798120763

Codutti, L., van Ingen, H., Vascotto, C., Fogolari, F., Corazza, A., Tell, G., Quadrifoglio, F., Viglino, P., Boelens, R., & Esposito, G. (2008). The solution structure of DNA-free Pax-8 paired box domain accounts for redox regulation of transcriptional activity in the pax protein family. *The Journal of Biological Chemistry*, *283*(48), 33321–33328. https://doi.org/10.1074/jbc.M805717200

Brickman, J. M., Jones, C. M., Clements, M., Smith, J. C., & Beddington, R. S. (2000). Hex is a transcriptional repressor that contributes to anterior identity and suppresses Spemann organiser function. *Development (Cambridge, England)*, *127*(11), 2303–2315. https://doi.org/10.1242/dev.127.11.2303

Bowden, S. A., & Goldis, M. (2024). *Congenital Hypothyroidism*.

Ossom Williamson, P., & Minter, C. I. J. (2019). Exploring PubMed as a reliable resource for scholarly communications services. *Journal of the Medical Library Association*, *107*(1). https://doi.org/10.5195/jmla.2019.433

Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., Pritzel, A., Wong, L. H., Zielinski, M., Sargeant, T., Schneider, R. G., Senior, A. W., Jumper, J., Hassabis, D., Kohli, P., & Avsec, Ž. (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*, *381*(6664). https://doi.org/10.1126/science.adg7492

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. https://doi.org/10.1093/nar/28.1.235

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Mooers, B. H. M. (2021). A PyMOL snippet library for Jupyter to boost researcher productivity. *Computing in Science & Engineering*, *23*(2), 47–53. https://doi.org/10.1109/mcse.2021.3059536

Yuan, S., Chan, H. C. S., Filipek, S., & Vogel, H. (2016). PyMOL and Inkscape Bridge the Data and the Data Visualization. *Structure*, *24*(12), 2041–2042. https://doi.org/10.1016/j.str.2016.11.012

Anaya, J.-M., Shoenfeld, Y., Rojas-Villarraga, A., Levy, R. A., & Cervera, R. (2013). *Autoimmunity: From Bench to Bedside [Internet]*

Pellizzari, L., Tell, G., & Damante, G. (1999). Co-operation between the PAI and RED subdomains of Pax-8 in the interaction with the thyroglobulin promoter. *The Biochemical Journal*, *337 (Pt 2)*(Pt 2), 253–262.

Congdon, T., Nguyen, L. Q., Nogueira, C. R., Habiby, R. L., Medeiros-Neto, G., & Kopp, P. (2001). A novel mutation (Q40P) in PAX8 associated with congenital hypothyroidism and thyroid hypoplasia: evidence for phenotypic variability in mother and child. *The*

Journal of Clinical Endocrinology and Metabolism, *86*(8), 3962–3967. https://doi.org/10.1210/jcem.86.8.7765

Vilain, C., Rydlewski, C., Duprez, L., Heinrichs, C., Abramowicz, M., Malvaux, P., Renneboog, B., Parma, J., Costagliola, S., & Vassart, G. (2001). Autosomal dominant transmission of congenital thyroid hypoplasia due to loss-of-function mutation of PAX8. *The Journal of Clinical Endocrinology and Metabolism*, *86*(1), 234–238. https://doi.org/10.1210/jcem.86.1.7140

Meeus, L., Gilbert, B., Rydlewski, C., Parma, J., Roussie, A. L., Abramowicz, M., Vilain, C., Christophe, D., Costagliola, S., & Vassart, G. (2004). Characterization of a novel loss of function mutation of PAX8 in a familial case of congenital hypothyroidism with in-place, normal-sized thyroid. *The Journal of Clinical Endocrinology and Metabolism*, *89*(9), 4285–4291. https://doi.org/10.1210/jc.2004-0166

Macchia, P. E., Lapi, P., Krude, H., Pirro, M. T., Missero, C., Chiovato, L., Souabni, A., Baserga, M., Tassi, V., Pinchera, A., Fenzi, G., Grüters, A., Busslinger, M., & di Lauro, R. (1998). PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. *Nature Genetics*, *19*(1), 83–86. https://doi.org/10.1038/ng0598-83

Carvalho, A., Hermanns, P., Rodrigues, A.-L., Sousa, I., Anselmo, J., Bikker, H., Cabral, R., Pereira-Duarte, C., Mota-Vieira, L., & Pohlenz, J. (2013). A new PAX8 mutation causing congenital hypothyroidism in three generations of a family is associated with abnormalities in the urogenital tract. *Thyroid : Official Journal of the American Thyroid Association*, *23*(9), 1074–1078. https://doi.org/10.1089/thy.2012.0649

Lanzerath, K., Bettendorf, M., Haag, C., Kneppo, C., Schulze, E., & Grulich-Henn, J. (2006). Screening for Pax8 mutations in patients with congenital hypothyroidism in South-West Germany. *Hormone Research*, *66*(2), 96–100. https://doi.org/10.1159/000093799

Bleu, M., Gaulis, S., Lopes, R., Sprouffske, K., Apfel, V., Holwerda, S., Pregnolato, M., Yildiz, U., Cordo′, V., Dost, A. F. M., Knehr, J., Carbone, W., Lohmann, F., Lin, C. Y., Bradner, J. E., Kauffmann, A., Tordella, L., Roma, G., & Galli, G. G. (2019). PAX8 activates metabolic genes via enhancer elements in Renal Cell Carcinoma. *Nature Communications*, *10*(1), 3739. https://doi.org/10.1038/s41467-019-11672-1

National Institutes of Health (US); Biological Sciences Curriculum Study. NIH Curriculum Supplement Series [Internet]. Bethesda (MD): National Institutes of Health (US); 2007. Understanding Human Genetic Variation.Available from: https://www.ncbi.nlm.nih.gov/books/NBK20363/#

Muzza, M., & Fugazzola, L. (2018). Thyroid gland development, molecular biology. *Encyclopedia of Endocrine Diseases*, 391–397. https://doi.org/10.1016/B978-0-12-801238-3.96023-9

Tordai, H., Torres, O., Csepi, M., Padányi, R., Lukács, G. L., & Hegedűs, T. (2024). Analysis of AlphaMissense data in different protein groups and structural context. *Scientific Data*, *11*(1), 495. https://doi.org/10.1038/s41597-024-03327-8

Iqbal, S., Pérez-Palma, E., Jespersen, J. B., May, P., Hoksza, D., Heyne, H. O., Ahmed, S. S., Rifat, Z. T., Rahman, M. S., Lage, K., Palotie, A., Cottrell, J. R., Wagner, F. F., Daly, M. J., Campbell, A. J., & Lal, D. (2020). Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants. *Proceedings*

of the National Academy of Sciences, *117*(45), 28201–28211. https://doi.org/10.1073/pnas.2002660117

Chang, L., Huang, R., Chen, J., Li, G., Shi, G., Xu, B., & Wang, L. (2023). An alpha-helix variant p.Arg156Pro in LMNA as a cause of hereditary dilated cardiomyopathy: genetics and bioinfomatics exploration. *BMC Medical Genomics*, *16*(1), 229. https://doi.org/10.1186/s12920-023-01661-1

Abrusán, G., & Marsh, J. A. (2016). Alpha Helices Are More Robust to Mutations than Beta Strands. *PLoS Computational Biology*, *12*(12), e1005242. https://doi.org/10.1371/journal.pcbi.10052

The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Appendix

All Found ClinVar Mutations Related to Hashimoto's Thyroiditis

Appendix Table 1

Variants Associated with Hashimoto's Thyroiditis based on ClinVar

DNA Base	Amino Acid Substitution	Gene	Type	Condition	Classification	Publications
c.106G	p.Asp36His	CEP128, TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Benign/Likely benign	Peeters RP et al, 2006
c.457G	p.Asp153Asn	CTLA4	Single nucleotide variant (missense variant)	Hashimoto thyroiditis	Conflicting classifications of pathogenicity	No publications found in the database
c.3416-1G		DUOX2	Single nucleotide variant (splice acceptor variant)	Congenital hypothyroidism	Likely pathogenic	No publications found in the database
c.3416-7C		DUOX2	Single nucleotide variant (intron variant)	Congenital hypothyroidism	Uncertain significance	No publications found in the database
c.2041+24G		IGSF1	Single nucleotide variant (intron variant)	X-linked central congenital hypothyroidism with late-onset	Benign	No publications found in the database
c.2042-1_2042del		IGSF1	Deletion (splice acceptor variant)	X-linked central congenital hypothyroidism with late-onset	Likely pathogenic	No publications found in the database
c.1614G	p.Trp538Ter	IGSF1	Single nucleotide variant (nonsense)	X-linked central congenital hypothyroidism with late-onset	Likely Pathogenic	No publications found in the database
c.3790C	p.Arg1264Ter	IGSF1	Single nucleotide variant (nonsense)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	No publications found in the database

c.3581dup	p.Glu1195fs	IGSF1	Duplication (frameshift variant)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Sun Y et al, 2012
c.3550C	p.Arg1184Ter	IGSF1	Single nucleotide variant (nonsense)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Nakamura A et al, 2013
c.2916G	p.Trp972Ter	IGSF1	Single nucleotide variant (nonsense)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Sun Y et al, 2012
c.2573C	p.Ser858Phe	IGSF1	Single nucleotide variant (missense variant)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Sun Y et al, 2012
c.2303T	p.Leu768Pro	IGSF1	Single nucleotide variant (missense variant)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Roche EF et al, 2018
c.2268dup	p.Arg757fs	IGSF1	Duplication (frameshift variant)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Tenenbaum-Rakover Y et al, 2016
c.2233del	p.Glu745fs	IGSF1	Deletion (frameshift variant)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Sun Y et al, 2012
c.2123_2149del	p.Ala708_Lys716del	IGSF1	Deletion (inframe_deletion)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	Sun Y et al, 2012
c.1814_1836dup	p.Arg613delinsGlyArgThrTer	IGSF1	Duplication (nonsense)	X-linked central congenital hypothyroidism with late-onset	Pathogenic	No publications found in the database
c.2407dup	p.His803fs	IGSF1	Duplication (frameshift variant)	X-linked central congenital hypothyroidism with late-onset	Pathogenic/Likely pathogenic	No publications found in the database

c.3813T	p.Val1271	IGSF1	Single nucleotide variant (synonymous variant)	X-linked central congenital hypothyroidism with late-onset	Uncertain significance	No publications found in the database
c.3669C	p.Asn1223Lys	IGSF1	Single nucleotide variant (missense variant)	X-linked central congenital hypothyroidism with late-onset	Uncertain significance	No publications found in the database
c.3467T	p.Val1156Glu	IGSF1	Single nucleotide variant (missense variant)	X-linked central congenital hypothyroidism with late-onset	Uncertain significance	No publications found in the database
c.3061G	p.Gly1021Arg	IGSF1	Single nucleotide variant (missense variant)	X-linked central congenital hypothyroidism with late-onset	Uncertain significance	No publications found in the database
c.2050C	p.Pro684Thr	IGSF1	Single nucleotide variant (missense variant)	X-linked central congenital hypothyroidism with late-onset	Uncertain significance	No publications found in the database
c.1672C	p.Gln558Ter	IGSF1	Single nucleotide variant (nonsense)	X-linked central congenital hypothyroidism with late-onset	Uncertain significance	No publications found in the database
c.2635C	p.His879Asp	IRS4	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.3161_3165del	p.Cys1054fs	IRS4	Deletion (frameshift variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Heinen CA et al, 2018
c.1772dup	p.Lys592fs	IRS4	Duplication (frameshift variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic/Likely pathogenic	Heinen CA et al, 2018
c.709G	p.Glu237Gln	LOC126806316, PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Likely benign	No publications found in the database
c.658C	p.Arg220Ter	LOC126806316, PAX8	Single nucleotide variant (nonsense)	Congenital hypothyroidism	Pathogenic	Fu C et al, 2015

c.777G	p.Gln259His	LOC126806316, PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.754C	p.Pro252Ala	LOC126806316, PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.602G	p.Ser201Asn	LOC126806316, PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.335-204del		NKX2-5	Deletion (3 prime UTR variant +2 more)	Hypothyroidism, congenital, nongoitrous	Likely pathogenic	No publications found in the database
c.*2478C		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.*2309A		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.*2146A		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.*1914C		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.*1044A		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.*1006A		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database

c.*1000T		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Benign	No publications found in the database
c.985T	p.Phe329Leu	PAX8	Single nucleotide variant (missense variant +2 more)	Hypothyroidism, congenital, nongoitrous	Benign/Likely benign	Carvalho A et al, 2013
c.*2276G		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Conflicting classifications of pathogenicity	No publications found in the database
c.1028A	p.Asn343Ser	PAX8	Single nucleotide variant (missense variant +1 more)	Hypothyroidism, congenital, nongoitrous	Conflicting classifications of pathogenicity	No publications found in the database
c.1009T	p.Ser337Ala	PAX8	Single nucleotide variant (missense variant +2 more)	Hypothyroidism, congenital, nongoitrous	Conflicting classifications of pathogenicity	Vizcaino MA et al, 2019
c.*627T		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Likely benign	No publications found in the database
c.1139C	p.Thr380Ile	PAX8	Single nucleotide variant (missense variant +1 more)	Hypothyroidism, congenital, nongoitrous	Likely benign	No publications found in the database
c.*2527A		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2517C		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2503_*2504dup		PAX8	Duplication (3 prime UTR variant)	Congenital hypothyroidism	Uncertain significance	No publications found in the database

c.*2475C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2409C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2368C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2326A	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2320G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2312G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2231C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2193G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2192C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*2044G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database

c.*2033C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1977T	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1942C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1807C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1765G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1490G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1457G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1393A	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1186G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1066G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database

c.*1060C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*1029G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*938C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*897C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*845G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*844C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*789G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*766C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*764C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*708C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database

c.*703G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*697C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*664T	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*612G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*527G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*501C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*472C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*423C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*372C	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*368G	PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database

c.*343G		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*290G		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.*204C		PAX8	Single nucleotide variant (3 prime UTR variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1277-10A		PAX8	Single nucleotide variant (intron variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1275G	p.Leu425	PAX8	Single nucleotide variant (synonymous variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1267A	p.Ser423Gly	PAX8	Single nucleotide variant (missense variant +1 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1226C	p.Pro409Arg	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1189G	p.Gly397Arg	PAX8	Single nucleotide variant (missense variant +2 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1185G	p.Val395	PAX8	Single nucleotide variant (synonymous variant +2 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1060C	p.Gln354Glu	PAX8	Single nucleotide variant (missense variant +2 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database

c.346G	p.Val116Ile	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Conflicting classifications of pathogenicity	No publications found in the database
c.215G	p.Arg72Gln	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Likely benign	No publications found in the database
c.237dup	p.Lys80fs	PAX8	Duplication (frameshift variant)	Hypothyroidism, congenital, nongoitrous	Likely pathogenic	No publications found in the database
c.236C	p.Ser79Phe	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Likely pathogenic	No publications found in the database
c.205G	p.Gly69Ser	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Likely pathogenic	No publications found in the database
c.203C	p.Thr68Ile	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Likely pathogenic	No publications found in the database
c.160A	p.Ser54Cys	PAX8,	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Likely pathogenic	No publications found in the database
c.397C	p.Arg133Trp	PAX8	Single nucleotide variant (missense variant)	Congenital hypothyroidism	Pathogenic	Vincenzi M et al, 2014
c.101T	p.Ile34Asn	PAX8	Single nucleotide variant (missense variant)	Congenital hypothyroidism	Pathogenic	Lanzerath K et al, 2006
c.119A	p.Gln40Pro	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism nongoitrous	Pathogenic	Congdon T et al, 2001

c.457_458del	p.Leu153fs	PAX8	Deletion (frameshift variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	No publications found in the database
c.322C	p.Arg108Ter	PAX8	Single nucleotide variant (nonsense)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Macchia PE et al, 1998
c.185T	p.Leu62Arg	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Macchia PE et al, 1998
c.170G	p.Cys57Tyr	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Vilain C et al, 2001
c.160A	p.Ser54Gly	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Meeus L et al, 2004
c.143C	p.Ser48Phe	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Grasberger H et al, 2005
c.92G	p.Arg31His	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Pathogenic	Macchia PE et al, 1998
c.297C	p.Asn99	PAX8	Single nucleotide variant (synonymous variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.858G	p.Gly286	PAX8	Single nucleotide variant (synonymous variant +1 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.817G	p.Asp273Asn	PAX8	Single nucleotide variant (missense variant +1 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database

c.898+12G		PAX8	Single nucleotide variant (non-coding transcript variant +1 more)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.324A	p.Arg108	PAX8	Single nucleotide variant (synonymous variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.280G	p.Asp94Asn	PAX8	Single nucleotide variant (missense variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.191+6C		PAX8	Single nucleotide variant (intron variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.441C	p.Cys147	PAX8	Single nucleotide variant (synonymous variant)	Hypothyroidism, congenital, nongoitrous	Uncertain significance	No publications found in the database
c.1076-1G		TG	Single nucleotide variant (splice acceptor variant)	Autoimmune thyroid disease	Likely pathogenic	Lo MS et al, 2018
c.7007G	p.Arg2336Gln	TG	Single nucleotide variant (missense variant)	Autoimmune thyroid disease	Likely pathogenic	Siffo S et al, 2018
c.2134dup	p.Ala712fs	TG	Duplication (frameshift variant)	Congenital hypothyroidism	Likely pathogenic	No publications found in the database
c.3452del	p.Val1151fs	TG	Deletion (frameshift variant)	Congenital hypothyroidism	Likely pathogenic	No publications found in the database
c.5182T	p.Cys1728Arg	TG	Single nucleotide variant (missense variant)	Congenital hypothyroidism	Pathogenic	No publications found in the database

c.-1623A		TG	Single nucleotide variant	Autoimmune thyroid disease	risk factor	Stefan M et al, 2011
c.6001G	p.Asp2001Asn	TG	Single nucleotide variant (missense variant)	Autoimmune thyroid disease	Uncertain significance	No publications found in the database
c.4481C	p.Pro1494Leu	TG	Single nucleotide variant (missense variant)	Autoimmune thyroid disease	Uncertain significance	No publications found in the database
c.6390T	p.Cys2130Trp	TG	Single nucleotide variant (missense variant)	Autoimmune thyroid disease	Uncertain significance	No publications found in the database
c.1566T	p.Ala522	TPO	Single nucleotide variant (synonymous variant)	Congenital hypothyroidism	Conflicting classifications of pathogenicity	No publications found in the database
c.1768+1_1768+9dup		TPO	Duplication (splice donor variant +1 more)	Congenital hypothyroidism	Likely pathogenic	No publications found in the database
c.866T	p.Phe289Ser	TPO	Single nucleotide variant (missense variant +1 more)	Congenital hypothyroidism	Pathogenic	No publications found in the database
c.650A	p.Asn217Ser	TPO	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Uncertain significance	No publications found in the database
c.*63dup		TPO	Duplication (3 prime UTR variant)	Congenital hypothyroidism	Uncertain significance	No publications found in the database
c.2161G	p.Val721Phe	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Conflicting classifications of pathogenicity	No publications found in the database
c.463A	p.Ile155Leu	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Conflicting classifications of pathogenicity	No publications found in the database

c.1342G	p.Val448Ile	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Conflicting classifications of pathogenicity	No publications found in the database
c.1170T	p.Cys390Trp	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Likely pathogenic	Beibermann H et al, 1997
c.1207G	p.Asp403Asn	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Likely pathogenic	No publications found in the database
c.1582C	p.Arg528Cys	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Likely pathogenic	Kara C et al, 2023
c.1777del	p.Ala593fs	TSHR	Deletion (frameshift variant)	Hypothyroidism due to TSH receptor mutations	Likely pathogenic	No publications found in the database
c.1960A	p.Ile654Phe	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Likely pathogenic	No publications found in the database
c.500T	p.Ile167Asn	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	Sunthornthepvarakul T et al, 1995
c.970C	p.Gln324Ter	TSHR	Single nucleotide variant (nonsense)	Hypothyroidism due to TSH receptor mutations	Pathogenic	de Roux N et al, 1996
c.1217_1234delinsCACG	p.Asn406fs	TSHR	Indel (frameshift variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	Beibermann H et al, 1997
c.1228G	p.Asp410Asn	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	de Roux N et al, 1996

c.1348del	p.Arg450fs	TSHR	Deletion (frameshift variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	Park KS, 2021
c.1400T	p.Leu467Pro	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	Alberti L et al, 2002
c.1430C	p.Thr477Ile	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	Tonacchera M et al, 2000
c.1575C	p.Phe525Leu	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	de Roux N et al, 1996
c.1798T	p.Cys600Arg	TSHR	Single nucleotide variant (missense variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic	Alberti L et al, 2002
c.1839C	p.Tyr613Ter	TSHR	Single nucleotide variant (nonsense)	Hypothyroidism due to TSH receptor mutations	Pathogenic	No publications found in the database
c.418del	p.Met140fs	TSHR	Deletion (frameshift variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic/Likely pathogenic	de Roux N et al, 1996
c.1963_1964del	p.Thr655fs	TSHR	Deletion (frameshift variant)	Hypothyroidism due to TSH receptor mutations	Pathogenic/Likely pathogenic	Calebiro D et al, 2012
c.1191G	p.Val397	TSHR	Single nucleotide variant (synonymous variant)	Hypothyroidism due to TSH receptor mutations	Uncertain significance	No publications found in the database
c.2475+2472T		ZFAT	Single nucleotide variant (intron variant)	Autoimmune thyroid disease	risk factor	Shirasawa S et al, 2004

c.1199G	p.Arg400Gln	ZFAT	Single nucleotide variant (missense variant +1 more)	Autoimmune thyroid disease	Uncertain significance	No publications found in the database