RJNMSR
Review Journal of Neurological & Medical Sciences Review

Volume 4, Issue 3, 2026

IN SILICO PREDICTION AND FUNCTIONAL ANALYSIS OF NONSYNONYMOUS SNPS OF HUMAN *TSHR* GENE

Javeria Nawaz[*1], Shameen Fatima[2], Shagufta Rasool[3], Areej Fatima[4], Farheena Iqbal Awan[5], Muhammad Usman Ghani[6]

[*1,2,3,4,5,6]*Centre for Applied Molecular Biology, University of the Punjab, Lahore, Pakistan*

[*1]Javeria3485@gmail.com

Corresponding Author: *
Javeria Nawaz

DOI: http://doi.org/10.5281/zenodo.19451501

Received	Accepted	Published
31 January 2026	15 March 2026	31 March 2026

ABSTRACT

TSHR (Thyroid-stimulating hormone receptor) is a vital thyrocyte membrane protein in the thyroid gland and a risk factor for TSH abnormalities. Genetic variations in the TSHR are directly involved in the pathogenesis of autoimmune hypothyroidism, Graves' disease, toxic adenomas, familial and sporadic non-autoimmune hypothyroidism, and forms of resistance to TSH. Identifying deleterious non-synonymous SNPs in the TSHR gene is crucial in understanding how these genetic variations affect its structure and function. In current study, potentially pathogenic nsSNPs were predicted using a combination of functional prediction tools, including SIFT, PROVEAN, PhD-SNP, SNP&GO and PolyPhen2. To further assess the impact of these nsSNPs on protein stability and function, MutPred, I-Mutant, Mupro and ConSurf, were employed. The 3D structure was predicted using TrRosetta, while gene and protein interaction network were explored using GeneMANIA and STRING, respectively. This analysis identified 5 pathogenic nsSNPs, rs142063461 (P68S), rs2075179 (N187K), rs121908871 (C390W), rs121908885 (L467), rs2140111252 (W488R) within the coding region of TSHR gene. The tools predicted decreasing in the protein stability and have higher RMSD values for these nsSNPs and showed deviation in structure from wild type TSHR protein. The prediction of gene-gene interaction showed the interaction of TSHR with other genes and its importance in different pathways, such as cAMP signaling pathway, MAPK, and GPCR pathway. The identified pathogenic nsSNPs of TSHR using computational analysis in this study could be used for large population-based investigations, diagnosis and can help in precision medicine.

Keywords: *In silico, non-synonymous SNPs, TSHR, Mutation, Deleterious, Prediction*

INTRODUCTION

Thyroid disease is an umbrella term composed of various conditions including hyperthyroidism, hypothyroidism, thyroid nodules, autoimmune thyroid disease, and thyroid cancers caused by disruption in the size, function and structure of the thyroid (1). This is the second most prevalent endocrine condition, following diabetes. An estimated 200 million people worldwide have been diagnosed with thyroid disorders (2).

Thyroid gland is the essential endocrine gland consist of follicular cells that are responsible for the synthesis, storage and secretion of thyroid hormones. These hormones play an important role in growth and development, cardiovascular function, carry normal cognition, bone health, metabolism and energy balance (3). The

production and secretion of the thyroid hormones are regulated by feedback loop involving the hypothalamus, pituitary gland and thyroid gland. The hypothalamus releases thyrotropin-releasing hormone (TRH) into the hypothalamic-hypophyseal portal system to the anterior pituitary gland. TRH stimulates thyrotropin cells in the anterior pituitary to secretes thyroid-stimulating hormones (TSH) into the blood where it binds to G-protein coupled thyrotropin receptor (thyroid Stimulating hormone receptor, TSHR) found on thyroid follicular cells and stimulation of the receptor release triiodothyronine 20% (T_3) and thyroxine 80% (T_4) hormones. The formation of (T_3) and (T_4) depend on the availability of iodide, stimulation by TSH stimulation, and the presence of tyrosine residues on thyroglobulin. Once (T_4) is released into the bloodstream, it can convert into (T_3) through the process of deiodination (4). The onset of thyroid disorders can be attributed to the imbalances production and secretion of thyroid hormones regulated by (5). Single chain thyrotropin receptor (TSHR) protein, found on the surface of thyroid follicular cells encoded by TSHR gene. TSHR gene is located on chromosome 14q31 covering almost 65 kb (kilo bases) of genomic DNA and consists of 10 exons code for a receptor protein of 764 amino acid residues (6). Mutations in the TSHR gene cause a mutant protein that does not bind to TSH or activate adenylate cyclase (7). Thus, mutant TSHR protein interferes with thyroid gland development and proper functioning. Mutations in the TSHR gene are inherited in an autosomal recessive manner, and exon 10 has been found to be a hotspot for mutations (8).

Human DNA exhibits various sequence variation among which SNPs (Single Nucleotide Polymorphism) is the most prevalent one, accounting for ~90% of genomic polymorphisms (9). Recent clinical and molecular studies have significantly advanced our understanding of nsSNPs in the TSHR gene and their association with thyroid disorders. A 2025 investigation of Iranian consanguineous families with congenital hypothyroidism (CH) identified multiple TSHR variants through exome sequencing, including the pathogenic missense variant rs121908863 (c.484C>G,p. Pro162Ala) in the leucine-rich repeat (LRR) domain, which disrupts TSH binding and cAMP signaling (10). Similarly, large-scale screening in Chinese populations has demonstrated the clinical significance of the rs189261858 (c.1349G> A, p. Arg450His) variant, a recurrent hotspot in exon 10 associated with TSH resistance and mild-to-moderate CH5. Functional characterization of these variants using in silico tools (SIFT, PolyPhen-2) and 3D protein modeling has clarified their structural impacts, such as destabilization of the TSH-binding ectodomain or disruption of G-protein coupling (11). Despite these advances, approximately 40% of TSHR-linked CH cases lack full mechanistic explanations, particularly regarding modifier genes and post-receptor signaling cascades (12). Building on this foundation, the current study employs advanced computational methods, including machine learning-based pathogenicity predictors, to systematically evaluate how nsSNPs alter TSHR gene functionally and structurally This approach aims to resolve unresolved questions about variant penetrance ultimately bridging the gap between genetic identification and functional understanding in thyroid pathophysiology.

Methodology:

This study employed a computational approach to evaluate the deleterious nsSNPs of TSHR gene, conducted in several steps using various bioinformatics tools is shown in Figure 1. GRCh38 was used as reference human genome for all tools and web servers.

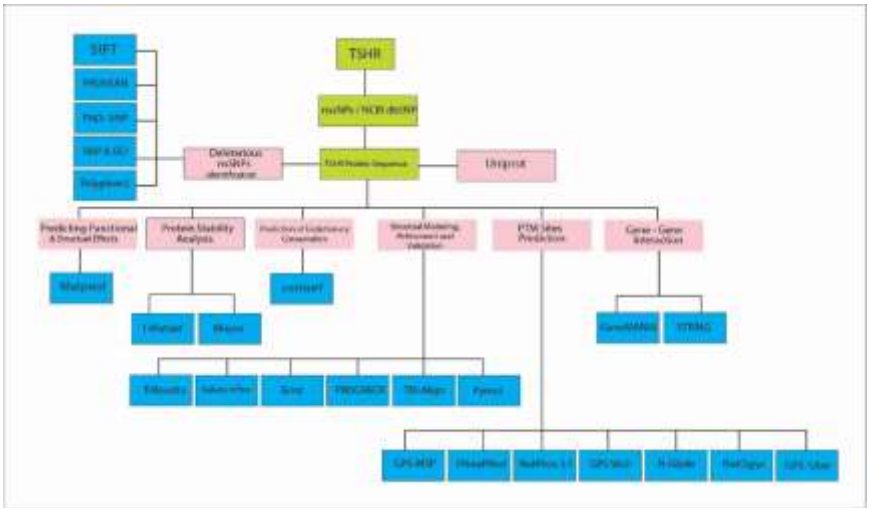


Figure 1 The figure represents the workflow of computational analysis of nsSNPs in TSHR gene

Data Mining:

TSHR gene nsSNPs were recruited from the NCBI dbSNP (gene ID: 7253) Database of (National Centre for Biotechnology Information) (https://www.ncbi.nlm.nih.gov/) (accessed on 6 January, 2025) (13). Gene View was selected of TSHR gene window in NCBI from where only missense SNPs that are reported in CLINVAR were selected for further work. These filtered nsSNPs were then validated with the 354 missense nsSNPs of ENSEMBL genome browser 113 (https://asia.ensembl.org/index.html) (accessed on 7 January, 2025) (14). The FASTA sequence of TSHR protein was obtained from the UniProt web server.

Screening of Deleterious SNPs:

Bioinformatics tools were employed to evaluate the effect of nsSNPs which were extracted after being filtered from multiple genomic databases mentioned previously. SIFT (Sorting Intolerant from Tolerant) (https://sift.bii.a-star.edu.sg/), a web-based tool that predicts the impact of nsSNPs on protein function based on sequence homology by assigning the score of ≤0.05, indicating a deleterious effect or a score of ≥0.05, indicating a tolerated variant (15-17). PROVEAN (Protein Variation Effect Analyzer) (http://provean.jcvi.org/seq_submit.php), another server was used for the prediction of changes in the biological function of a protein caused by an amino acid substitution. It gives score on the basis of degree of conservation across species with a cut-off value of -2.5. A score below -2.5 was considered deleterious, while a score above -2.5 was considered neutral (18). PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms) (https://snps.biofold.org/phd-snp/phd-snp.html), a tool which predicts the pathogenicity of nsSNPs, machine learning models, providing a score to classify them as disease-causing or neutral. It gives reliability index of nsSNPs ranging from 0-10 with 0 being the lowest and 10 being the highest confidence score (19-21). SNPs&GO (https://snps.biofold.org/snps-and-go/snps-and-go.html), based on the information of molecular and functional information from the gene ontology (GO) database, calculates deleterious effect of nsSNPs, where resultant output provides a probability value. Probability values greater than 0.5 are considered to represent nsSNPs as disease-causing (22-25). PolyPhen-2 (Polymorphism Phenotyping 2) (http://genetics.bwh.harvard.edu/pph2/), which classifies nsSNPs as benign, probably damaging and possibly damaging. The cut-off value (score) for benign is <0.5, possibly damaging varies from 0.5–0.8, and for probably damaging is ≥0.9 (19, 23, 26, 27). Together these tools provided a comprehensive evaluation of nsSNPs, enabling the identification of potentially harmful variants.

Analysis of Structural and Functional Variants:
The structural and functional impact of nsSNPs, resulting from amino acid change was predicted using MutPred2 (http://mutpred.mutdb.org/), a web-based tool which analyzes amino acid substitution and predicts their possible molecular role in disease. The FASTA sequence of TSHR protein was submitted and deleterious nsSNPs were selected. A score greater than 0.5 indicates a mutation that may be pathogenic (9).

Evaluation of Protein Stability:
I-Mutant 2.0 (https://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) (28) was used to predict the protein stability. This web server is based upon support vector machines (SVMs) algorithm to predict mutated protein stability with corresponding ΔΔG (change in Gibbs free energy) values. The result shows in three sections, prediction section, reliability index (RI) and a DDG (Delta Delta G) values, which is determined by subtracting the wild type unfolding Gibbs free energy value (Kcal/mol) from the mutant type unfolding Gibbs free energy value (29). The RI of the nsSNPs ranging from 0-10, where 0 shows lowest reliability and 10 shows highest reliability. The effect of deleterious nsSNPs was predicted by submitting protein sequence of TSHR protein having conditions set at 25°C temperature and 7.0 pH. MUpro tool (http://mupro.proteomics.ics.uci.edu/) is also used to analyze change in protein stability due to nsSNPs. The sever composed of two software machine learning methods: neural network and SVM. The protein calculates the confidence of prediction, which is between -1 and 1. When the protein stability decreases due to mutation it is reported as confidence score <0 and when the confidence score is > 0 the mutation increases the protein stability (30).

Evolutionary Conservation Analysis of Protein:
The degree of conservation of an amino acid position shows its structural and functional importance. The evolutionary conservation of each amino acid present in the protein sequence was evaluated using ConSurf (https://consurf.tau.ac.il/consurf_index.php). This computational tool compared the query protein sequence with related sequences across multiple species (31).

Structural Modelling, Refinement and Validation:
3D Protein Structure Prediction:
Firstly, Phyre2 (https://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) tool was used for prediction of 3D structure. Then, TrRosetta (https://yanglab.qd.sdu.edu.cn/trRosetta/help/) was also used to predict protein 3D structure using FASTA sequence. It predicts inter-residue orientations and distances from multiple sequence alignments (MSA) using deep learning. A quick gradient-based energy minimization approach is then used to put these anticipated geometric restrictions together into 3D structures (32).

Refinement of Structure:
The GalaxyRefine (https://galaxy.seoklab.org/refine) web server used for protein structure refinement. This server is based on a refinement method that performs short molecular dynamics(MD) relaxations after repeated side chain repacking perturbations (33).
Validation:
The 3D protein model was further assessed using the Structure Validation Server v6.1 (https://saves.mbi.ucla.edu/).It uses multiple tools to assess the quality and reliability of predicted 3D structures (34).

Template Modelling Alignment:
The TM-align (https://zhanggroup.org/TM-align/) performs a residual-residual alignment to compare wild and mutant structure. The TM-score, which ranges between 0 to 1, and the RMSD (Root Mean Square Deviation), which measures the difference between the wild and mutant structure. The RMSD shows the amount of variation or divergence between the two structures, and the TM-score shows the structural similarity (1 denotes a perfect match) (30). PyMol (https://pymol.org/) is evaluated for its ability to visualizing, analyzing, and manipulating (35).

PTM Sites Predictions:

Protein Function can be predicted through a thorough study on post-translational modification (PTM) sites in protein. Methylation sites in TSHR protein were predicted by GPS-MSP 3.0 (http://msp.biocuckoo.org/online.php) (36) and PRmePRed (https://bioinfo.icgeb.res.in/PRmePRed/svmtool4.php) (37). At serine, tyrosine and threonine residual positions at TSHR protein sequence, prediction of phosphorylation sites was done using NetPhos3.1 (https://services.healthtech.dtu.dk/services/NetPhos-3.1/) and (https://gps.biocuckoo.cn/online.php). In comparison to NetPhos 3.1, GPS 3.0 predicted more specific results with a greater chance of phosphorylation. NetPhos 3.1 made use of neural network ensembles and a 0.5 cutoff was threshold (38). Residues having higher score than this threshold were predicted to be phosphorylated. For prediction of Ubiquitylation sites in TSHR protein, GPS-Uber (https://gpsuber.biocuckoo.cn/online.php) was used (39). Another important post-translation modification event in TSHR protein is glycosylation which was predicted by using N-Glyde 1.0 (https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/) and NetOglyc 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/). In N-Glyde 1.0 , a score more than 0.6 of residues possess to be glycosylated (40, 41).

Gene-Gene Interaction of TSHR:

Interaction of TSHR gene was studied through GeneMANIA (https://genemania.org/) and STRING (https://string-db.org/) to observe its association with other genes and to predict the effects of nsSNPs of TSHR on other related gene (42). GeneMANIA employs several parameters such as genetic and protein interaction, co-localization, co-expression, pathways and proteins domain similarities to predict how a given gene will interact with several additional genes (43). The STRING creates a Protein-Protein Interaction (PPI) network by connecting known proteins to other proteins either directly and indirectly. PPI play a Vital role in analyzing the functional connections of all proteins in the cell (44).

Results:

Recruited nsSNPs:

A total of 80237 SNPs were recruited from dbSNP, that is the largest database for SNPs which consisted of 440 synonymous SNPs, 1052 missense SNPs, 309 located in 5' UTR, 677 occurred in 3'UTR and 73105 were in intronic region is given in Figure 2. Only 81 nsSNPs of CLINVAR were selected for further analysis.

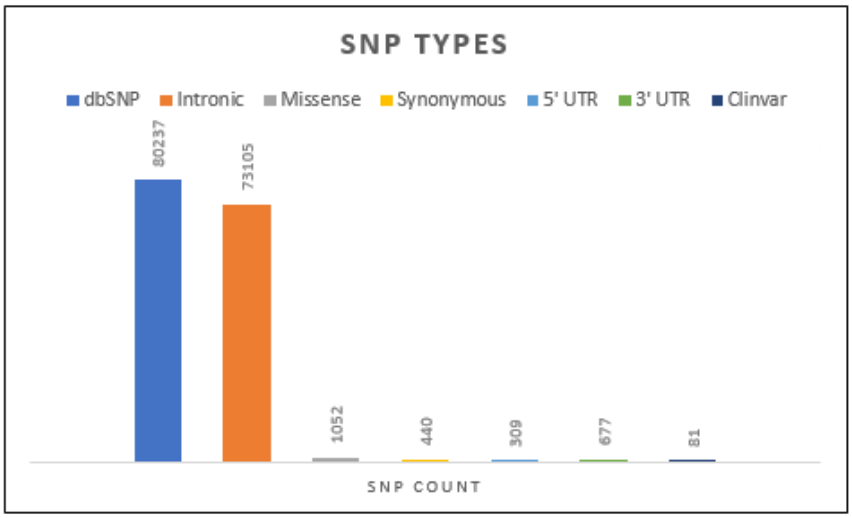


Figure 2 Distribution of SNPs in TSHR gene, including missense, synonymous, intronic variant, 5' UTR, 3' UTR and nsSNPs in ClinVar

Identification of deleterious nsSNPs:

All nsSNPs were analyzed using different bioinformatic alogrithm consist of SIFT, PROVEAN, PhD-SNP and SNP&GO. In SIFT, among the 81 nsSNPs, 47 nsSNPs were found to be deleterious having SIFT score below 0.05. In PROVEAN, 42 nsSNPs were found as deleterious having a score below -2.5, a cut-off value of this tool. PhD-SNP identified 45 nsSNPs to be disease-causing. SNP&GO also predicted the 62 SNPs from previous tool as disease-causing. Results of all tools are given in Table 1 and Figure 3. These filtered nsSNPs were then submitted to Polyphen2 software, among which the 32 probably damaging, 1 benign and 2 possible damaging nsSNPs. Results from polyphen2 tools are given in Table 2.

Table 1 The Results of The Four In-Silico Tools (SIFT, PROVEAN, PhD-SNP And SNP&GO) For All Filtered nsSNPs of TSHR Gene

Sr.	SNP ID	A.A change	SIFT		Provean		PHD-SNP			SNP & GO		
			Score	Prediction	Score	Prediction	Prediction	RI	Probability	Prediction	RI	Probability
1	rs45499704	E34K	0.92	Tolerated	0.153	Neutral	Neutral	5	0.236	Disease	2	0.608
2	rs61747482	D36H	0.508	Tolerated	0.104	Neutral	Neutral	8	0.254	Disease	0	0.52
3	rs121908869	C41S	0.099	Tolerated	-2.756	Deleterious	Disease	6	0.72	Disease	8	0.917
4	rs147137913	S49G	0.357	Tolerated	-1.204	Neutral	Neutral	6	0.105	Neutral	6	0.914
5	rs2234919	P52T	0.585	Tolerated	0.489	Neutral	Neutral	7	0.08	Neutral	2	0.416
6	rs886050853	S53R	0.13	Tolerated	-0.536	Neutral	Neutral	3	0.398	Disease	2	0.583
7	rs781625203	T56I	0.32	Tolerated	-1.052	Neutral	Neutral	7	0.228	Disease	1	0.547
8	rs200401152	L57P	0.062	Tolerated	-3.156	Deleterious	Disease	6	0.665	Disease	6	0.81
9	rs1886292569	T66N	0.04	Deleterious	-1.792	Neutral	Neutral	4	0.452	Disease	6	0.798
10	rs142063461	P68S	0.013	Deleterious	-4.078	Deleterious	Disease	1	0.543	Disease	7	0.836
11	rs121908865	R109Q	0.757	Tolerated	-0.282	Neutral	Neutral	6	0.198	Disease	2	0.61
12	rs760874290	G132R	0	Deleterious	-1.896	Neutral	Disease	4	0.62	Disease	7	0.838
13	rs1888697548	T136A	0	Deleterious	-3.734	Deleterious	Disease	6	0.529	Disease	8	0.887
14	rs141293178	I155L	0.607	Tolerated	-0.202	Neutral	Neutral	2	0.408	Disease	1	0.564
15	rs121908863	P162S	0.495	Tolerated	-1.19	Neutral	Neutral	8	0.246	Neutral	4	0.301
16	rs121908863	P162A	0.495	Tolerated	-0.746	Neutral	Neutral	6	0.207	Neutral	5	0.243
17	rs121908862	I167N	0	Deleterious	-5.422	Deleterious	Disease	9	0.906	Disease	9	0.954

18	rs121908879	K183R	0.67	Tolerated	-0.059	Neutral	Neutral	5	0.156	Neutral	4	0.309
19	rs2075179	N187K	0	Deleterious	-4.969	Deleterious	Disease	9	0.908	Disease	9	0.96
20	rs760702366	A204V	0.06	Tolerated	-1.623	Neutral	Neutral	6	0.333	Disease	3	0.653
21	rs144127341	V233M	0.032	Deleterious	-1.684	Neutral	Neutral	8	0.257	Neutral	0	0.494
22	rs189506473	G245S	0.025	Deleterious	-4.736	Deleterious	Disease	7	0.752	Disease	7	0.834
23	rs1891604207	I253T	0.6	Tolerated	-0.207	Neutral	Neutral	6	0.375	Disease	3	0.642
24	rs121908873	S281N	0	Deleterious	-2.496	Neutral	Disease	5	0.679	Disease	8	0.906
25	rs121908873	S281I	0	Deleterious	-5.011	Deleterious	Disease	7	0.747	Disease	8	0.916
26	rs756847429	E303V	0.3	Tolerated	-0.243	Neutral	Disease	3	0.309	Disease	1	0.57
27	rs142122217	S305R	0.533	Tolerated	-0.904	Neutral	Neutral	7	0.232	Neutral	6	0.224
28	rs121908882	R310C	0.085	Tolerated	-0.779	Neutral	Disease	2	0.591	Neutral	4	0.316
29	rs139286618	R310H	0.239	Tolerated	-0.443	Neutral	Neutral	3	0.298	Neutral	5	0.231
30	rs121908867	Q324K	0.24	Tolerated	-0.639	Neutral	Neutral	8	0.146	Neutral	6	0.215
31	rs376188008	D345N	0.08	Tolerated	-1.043	Neutral	Neutral	3	0.505	Disease	0	0.525
32	rs370857348	N349K	0.21	Tolerated	-1.37	Neutral	Neutral	6	0.39	Disease	3	0.631
33	rs121908871	C390W	0	Deleterious	-6.696	Deleterious	Disease	9	0.958	Disease	9	0.967
34	rs199702292	C408R	0	Deleterious	-10.15	Deleterious	Disease	9	0.94	Disease	9	0.971
35	rs121908868	D410N	0.004	Deleterious	-4.227	Deleterious	Disease	5	0.67	Disease	8	0.917
36	rs587778742	V424I	0.98	Tolerated	0.112	Neutral	Neutral	9	0.19	Neutral	2	0.399
37	rs121908883	G431S	0	Deleterious	-4.966	Deleterious	Disease	6	0.904	Disease	8	0.911
38	rs2140110634	F434S	0.03	Deleterious	-1.239	Neutral	Neutral	4	0.493	Disease	6	0.808
39	rs777698828	N447K	0.02	Deleterious	-0.948	Neutral	Neutral	2	0.374	Neutral	2	0.422
40	rs201889708	V448I	0.135	Tolerated	-0.795	Neutral	Neutral	4	0.485	Disease	4	0.714
41	rs189261858	R450H	0.001	Deleterious	-4.476	Deleterious	Disease	7	0.732	Disease	8	0.902

42	rs121908864	M453T	0.001	Deleterious	-5.166	Deleterious	Disease	7	0.824	Disease	9	0.959
43	rs121908885	L467P	0	Deleterious	-6.413	Deleterious	Disease	8	0.931	Disease	9	0.935
44	rs1595178231	T477P	0	Deleterious	-4.918	Deleterious	Disease	9	0.912	Disease	9	0.959
45	rs121908881	T477I	0.001	Deleterious	-5.038	Deleterious	Disease	8	0.788	Disease	9	0.945
46	rs2140111252	W488R	0	Deleterious	-12.36	Deleterious	Disease	8	0.946	Disease	9	0.958
47	rs121908876	S505N	0.005	Deleterious	-2.285	Neutral	Disease	8	0.858	Disease	7	0.841
48	rs772490052	S508L	0	Deleterious	-5.197	Deleterious	Disease	9	0.948	Disease	9	0.96
49	rs121908874	V509A	0.001	Deleterious	-3.465	Deleterious	Disease	8	0.773	Disease	5	0.758
50	rs780018604	R519H	0	Deleterious	-4.504	Deleterious	Disease	7	0.918	Disease	9	0.947
51	rs200138601	F525S	0.179	Tolerated	-3.475	Deleterious	Disease	7	0.845	Disease	9	0.942
52	rs121908870	F525L	0.29	Tolerated	-3.385	Deleterious	Disease	5	0.793	Disease	7	0.832
53	rs571893270	R528H	0.05	Tolerated	-1.51	Neutral	Neutral	4	0.48	Disease	6	0.825
54	rs150602845	R534C	0.002	Deleterious	-5.895	Deleterious	Disease	6	0.823	Disease	8	0.883
55	rs747680932	I541V	0.46	Tolerated	-0.413	Neutral	Neutral	8	0.245	Neutral	3	0.334
56	rs121908872	A553T	0.002	Deleterious	-3.457	Deleterious	Disease	6	0.715	Disease	7	0.87
57	rs1237819854	L557W	0	Deleterious	-4.917	Deleterious	Disease	5	0.795	Disease	7	0.825
58	rs2140112304	M572R	0	Deleterious	-5.352	Deleterious	Disease	5	0.803	Disease	9	0.957
59	rs61742289	T574S	0.003	Deleterious	-1.899	Neutral	Neutral	6	0.252	Neutral	2	0.417
60	rs751266534	Y582F	0	Deleterious	-3.1	Deleterious	Neutral	3	0.766	Disease	5	0.766
61	rs142632518	V595I	0.735	Tolerated	0.144	Neutral	Neutral	8	0.08	Neutral	7	0.151
62	rs121908884	C600R	0.001	Deleterious	-10.62	Deleterious	Disease	8	0.944	Disease	9	0.95
63	rs121908859	D619G	0.004	Deleterious	-6.146	Deleterious	Disease	8	0.855	Disease	6	0.804
64	rs121908877	L629F	0	Deleterious	-3.407	Deleterious	Disease	6	0.825	Disease	7	0.848
65	rs121908861	F631L	0	Deleterious	-5.336	Deleterious	Disease	6	0.785	Disease	5	0.762

66	rs28937584	D633H	0.001	Deleterious	-5.754	Deleterious	Disease	9	0.932	Disease	8	0.916
67	rs1891823429	C636G	0	Deleterious	-10.67	Deleterious	Disease	9	0.968	Disease	10	0.978
68	rs121908880	P639S	0	Deleterious	-7.111	Deleterious	Disease	7	0.921	Disease	9	0.941
69	rs587778741	Y643C	0	Deleterious	-5.479	Deleterious	Neutral	0	0.666	Disease	5	0.767
70	rs2140113209	A644V	0	Deleterious	-3.521	Deleterious	Disease	5	0.66	Disease	3	0.656
71	rs761428348	P652A	0	Deleterious	-6.975	Deleterious	Neutral	1	0.279	Neutral	5	0.269
72	rs767239688	I654F	0	Deleterious	-3.455	Deleterious	Disease	6	0.766	Disease	3	0.671
73	rs143773384	S657R	0	Deleterious	-3.22	Deleterious	Disease	5	0.687	Disease	3	0.664
74	rs2140113502	F666S	0	Deleterious	-6.592	Deleterious	Disease	7	0.746	Disease	3	0.668
75	rs121908875	C672Y	0.006	Deleterious	-7.656	Deleterious	Disease	9	0.948	Disease	9	0.947
76	rs368452281	R707Q	0.016	Deleterious	-2.445	Neutral	Disease	7	0.737	Disease	4	0.708
77	rs617454049	V721F	0.522	Tolerated	-0.091	Neutral	Neutral	6	0.263	Neutral	6	0.205
78	rs1991517	E727D	0.756	Tolerated	-0.018	Neutral	Neutral	9	0.117	Neutral	7	0.15
79	rs61743974	N744K	0.415	Tolerated	-0.018	Neutral	Neutral	6	0.64	Disease	4	0.682
80	rs763398101	G753S	0.34	Tolerated	0.202	Neutral	Neutral	6	0.572	Neutral	0	0.488
81	rs1891848268	Q761P	0.1	Tolerated	0.028	Neutral	Neutral	6	0.646	Disease	5	0.73

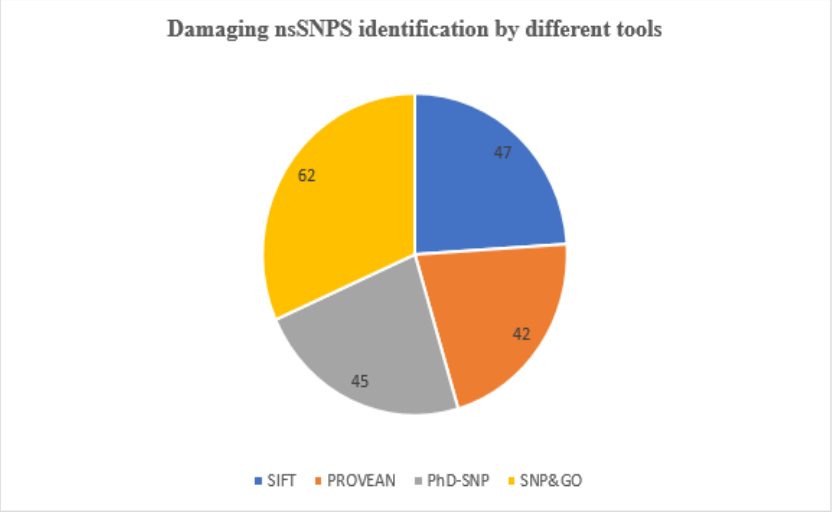


Figure 3 Number of potential damaging nsSNPs identification by different computational tools.

Table 2 Polyphen2 Analysis of nsSNPs In TSHR

Sr. No.	SNP	A.Amino change	Polyphen-2	
			Prediction	Score
1	rs142063461	P68S	Probably Damaging	0.939
2	rs1888697548	T136A	Benign	0.274
3	rs121908862	I167N	Probably Damaging	0.986
4	rs2075179	N187K	Probably Damaging	0.999
5	rs189506473	G245S	Probably Damaging	0.995
6	rs121908873	S281I	Possibly Damaging	0.893
7	rs121908871	C390W	Probably Damaging	0.999
8	rs199702292	C408R	Probably Damaging	0.999
9	rs121908868	D410N	Probably Damaging	1
10	rs121908883	G431S	Probably Damaging	1
11	s189261858	R450H	Probably Damaging	1
12	rs121908864	M453T	Probably Damaging	0.997
13	rs121908885	L467P	Probably Damaging	1
14	rs1595178231	T477P	Probably Damaging	1
15	rs121908881	T477I	Probably Damaging	1
16	rs2140111252	W488R	Probably Damaging	1
17	rs772490052	S508L	Probably Damaging	0.999
18	rs121908874	V509A	Probably Damaging	0.999
19	rs780018604	R519H	Probably Damaging	1
20	rs150602845	R534C	Probably Damaging	0.956
21	rs121908872	A553T	Probably Damaging	0.987
22	rs1237819854	L557W	Probably Damaging	0.974
23	rs2140112304	M572R	Probably Damaging	0.998
24	rs121908884	C600R	Probably Damaging	1

25	rs121908859	D619G	Probably Damaging	0.992
26	rs121908877	L629F	Probably Damaging	0.928
27	rs121908861	F631L	Probably Damaging	0.998
28	rs28937584	D633H	Probably Damaging	0.998
29	rs1891823429	C636G	Probably Damaging	0.999
30	rs121908880	P639S	Probably Damaging	1
31	rs2140113209	A644V	Probably Damaging	1
32	rs767239688	I654F	Probably Damaging	0.999
33	rs143773384	S657R	Possibly Damaging	0.895
34	rs2140113502	F666S	Probably Damaging	0.999
35	rs121908875	C672Y	Probably Damaging	1

Prediction of Structural and Functional Impact using MutPred:

All the filtered 35 Polyphen2 nsSNPs were taken to MutPred tool, it gave results with probability score which are given in Table 3. This tool predicted the pathogenicity of nsSNPs including altered transmembrane protein, altered ordered interface, loss/gain of allosteric site, gain of N-linked glycosylation, gain of catalytic site, altered metal binding, gain of relative solvent accessibility, gain of GPI-anchor amidation and loss/gain of strand, showing that many nsSNPs can cause protein alterations. Information for the mentioned features along with ***p***-values is provided in Table 4.

Table 3 Effects of nsSNPs on structural and functional properties of TSHR by MutPred sever

Sr.	Mutation	Probability of deleterious mutation	Molecular mechanisms with P-values <= 0.05	Affected PROSITE and ELM Motifs
1	P68S	0.672	Gain of Strand (Pr = 0.27 | P = 0.02) Altered Transmembrane protein (Pr = 0.12 | P = 0.03)	ELME000173 ELME000197 ELME000334
2	T136A	0.58	Loss of Strand (Pr = 0.26 | P = 0.04) Altered Transmembrane protein (Pr = 0.15 | P = 0.01)	PS00008
3	I167N	0.871	Altered Transmembrane protein (Pr = 0.28 | P = 4.6e-04) Gain of Sulfation at Y163 (Pr = 0.02 | P = 0.04)	None
4	N187K	0.871	Altered Transmembrane protein (Pr = 0.27 | P = 7.2e-04) Gain of Loop (Pr = 0.27 | P = 0.03) Gain of Strand (Pr = 0.26 | P = 0.04) Altered Stability (Pr = 0.14 | P = 0.02)	ELME000084 ELME000122 ELME000146
5	G245S	0.695	Gain of Intrinsic disorder (Pr = 0.30 | P = 0.04) Altered Transmembrane protein (Pr = 0.17 | P = 9.3e-03)	None
6	S281I	0.857	Altered Transmembrane protein (Pr = 0.15 | P = 0.01)	ELME000063
7	C390W	0.791	Loss of Disulfide linkage at C390 (Pr = 0.33 | P = 8.1e-04)	

			Gain of Strand (Pr = 0.28 | P = 7.6e-03) Altered Transmembrane protein (Pr = 0.27 | P = 7.1e-04) Loss of Sulfation at Y385 (Pr = 0.23 | P = 6.2e-04) Gain of Proteolytic cleavage at D386 (Pr = 0.15 | P = 0.01)	ELME000147 ELME000182
8	C408R	0.966	Gain of Intrinsic disorder (Pr = 0.33 | P = 0.03) Altered Metal binding (Pr = 0.28 | P = 0.02) Altered Transmembrane protein (Pr = 0.27 | P = 6.0e-04) Loss of Disulfide linkage at C408 (Pr = 0.24 | P = 7.7e-03)	None
9	D410N	0.686	Altered Transmembrane protein (Pr = 0.28 | P = 4.0e-04) Altered Ordered interface (Pr = 0.25 | P = 0.02) Altered Metal binding (Pr = 0.24 | P = 0.04) Gain of Disulfide linkage at C408 (Pr = 0.21 | P = 0.02)	None
10	G431S	0.816	Loss of Strand (Pr = 0.29 | P = 4.1e-03) Gain of Helix (Pr = 0.28 | P = 0.02) Altered Transmembrane protein (Pr = 0.03 | P = 1.6e-03) Gain of GPI-anchor amidation at N432 (Pr = 0.01 | P = 0.03)	ELME000045 ELME000197 ELME000333
11	R450H	0.62	Altered Ordered interface (Pr = 0.26 | P = 0.01)	ELME000106
12	M453T	0.918		ELME000052 ELME000106 ELME000122 ELME000146
13	L467P	0.953	Altered Ordered interface (Pr = 0.26 | P = 0.01) Altered Transmembrane protein (Pr = 0.13 | P = 0.02)	ELME000120 ELME000182 ELME000294
14	T477P	0.861	Altered Ordered interface (Pr = 0.33 | P = 8.0e-03) Loss of Helix (Pr = 0.28 | P = 0.02) Altered Transmembrane protein (Pr = 0.27 | P = 6.5e-04) Loss of Relative solvent accessibility (Pr = 0.26 | P = 0.03) Gain of Allosteric site at D474 (Pr = 0.21 | P = 0.03) Altered Metal binding (Pr = 0.16 | P = 0.03)	ELME000064 ELME000182 ELME000220 PS00006
15	T477I	0.797	Altered Ordered interface (Pr = 0.30 | P = 0.02) Altered Transmembrane protein (Pr = 0.28 | P = 4.6e-04) Loss of Relative solvent accessibility (Pr = 0.26 | P = 0.03)	ELME000064 ELME000081 ELME000182 ELME000220 PS00006
16	W488R	0.956	Gain of Relative solvent accessibility (Pr = 0.33 | P = 3.0e-03) Altered Ordered interface (Pr = 0.31 | P = 0.01) Altered Metal binding (Pr = 0.29 | P = 5.5e-03)	ELME000062 ELME000097

			Loss of Allosteric site at W488 (Pr = 0.29 | P = 6.1e-03) Gain of Helix (Pr = 0.28 | P = 0.02) Gain of Catalytic site at W488 (Pr = 0.18 | P = 0.01) Gain of Pyrrolidone carboxylic acid at Q489 (Pr = 0.12 | P = 7.2e-03)	
17	S508L	0.815	Gain of Helix (Pr = 0.27 | P = 0.04) Altered Ordered interface (Pr = 0.24 | P = 0.03) Altered Transmembrane protein (Pr = 0.20 | P = 5.2e-03)	ELME000063 ELME000147 ELME000335 ELME000336 PS00237
18	V509A	0.57	Altered Ordered interface (Pr = 0.26 | P = 0.02) Altered Transmembrane protein (Pr = 0.19 | P = 7.0e-03)	ELME000052 ELME000063 ELME000085 ELME000147 PS00237
19	R519H	0.837	Loss of Helix (Pr = 0.27 | P = 0.04) Gain of Strand (Pr = 0.27 | P = 0.02) Altered Metal binding (Pr = 0.15 | P = 0.04)	ELME000302 PS00237
20	R534C	0.382	-	-
21	A553T	0.759	Altered Transmembrane protein (Pr = 0.03 | P = 2.3e-03)	ELME000335 ELME000336
22	L557W	0.848	Altered Transmembrane protein (Pr = 0.02 | P = 4.4e-03)	ELME000333
23	M572R	0.923	Altered Transmembrane protein (Pr = 0.29 | P = 2.6e-04) Altered Signal peptide (Pr = 0.01 | P = 5.5e-03)	ELME000062
24	C600R	0.925	Altered Ordered interface (Pr = 0.27 | P = 7.1e-03) Altered Transmembrane protein (Pr = 0.01 | P = 0.01)	ELME000146 ELME000233
25	D619G	0.873	Loss of Relative solvent accessibility (Pr = 0.27 | P = 0.03) Altered Transmembrane protein (Pr = 0.22 | P = 2.8e-03) Loss of Allosteric site at K624 (Pr = 0.19 | P = 0.05) Altered DNA binding (Pr = 0.15 | P = 0.04) Gain of Methylation at K624 (Pr = 0.09 | P = 0.05) Loss of GPI-anchor amidation at N614 (Pr = 0.05 | P = 3.9e-03)	ELME000233 PS00008
26	L629F	0.934	Loss of Relative solvent accessibility (Pr = 0.29 | P = 0.01) Gain of Strand (Pr = 0.27 | P = 0.02) Gain of Allosteric site at R625 (Pr = 0.26 | P = 8.3e-03) Altered Transmembrane protein (Pr = 0.21 | P = 4.2e-03) Altered DNA binding (Pr = 0.18 | P = 0.03) Altered Metal binding (Pr = 0.13 | P = 6.7e-03) Loss of Catalytic site at D633 (Pr = 0.12 | P = 0.03)	ELME000149

27	F631L	0.931	Loss of Strand (Pr = 0.28 | P = 0.01) Gain of Helix (Pr = 0.27 | P = 0.04) Altered Transmembrane protein (Pr = 0.16 | P = 0.01) Altered Metal binding (Pr = 0.13 | P = 5.6e-03) Loss of Catalytic site at D633 (Pr = 0.13 | P = 0.03)	ELME000052 ELME000333
28	D633H	0.947	Altered Metal binding (Pr = 0.31 | P = 2.6e-04) Gain of Strand (Pr = 0.27 | P = 0.02) Altered Transmembrane protein (Pr = 0.13 | P = 0.02) Loss of Catalytic site at D633 (Pr = 0.12 | P = 0.03)	ELME000052 ELME000336
29	C636G	0.945	Altered Metal binding (Pr = 0.21 | P = 0.03) Gain of Catalytic site at D633 (Pr = 0.15 | P = 0.02) Altered Transmembrane protein (Pr = 0.03 | P = 2.0e-03)	ELME000052
30	P639S	0.901	Altered Ordered interface (Pr = 0.24 | P = 0.04) Altered Transmembrane protein (Pr = 0.03 | P = 1.3e-03)	ELME000197 ELME000337
31	A644V	0.577	Altered Transmembrane protein (Pr = 0.14 | P = 0.02)	ELME000182
32	I654F	0.898	Altered Transmembrane protein (Pr = 0.30 | P = 1.3e-04) Loss of GPI-anchor amidation at N658 (Pr = 0.01 | P = 0.03)	ELME000053 ELME000106 ELME000146 ELME000237 ELME000333 ELME000336
33	S657R	0.716	Altered Transmembrane protein (Pr = 0.27 | P = 6.9e-04) Gain of Loop (Pr = 0.27 | P = 0.04) Loss of GPI-anchor amidation at N658 (Pr = 0.01 | P = 0.03)	ELME000053 ELME000062 ELME000233 ELME000237 ELME000333 ELME000337 PS00005
34	F666S	0.92	Altered Transmembrane protein (Pr = 0.34 | P = 2.4e-05) Altered Stability (Pr = 0.16 | P = 0.02) Loss of Sulfation at Y667 (Pr = 0.04 | P = 0.01)	ELME000167 ELME000333 ELME000336
35	C672Y	0.937	Altered Transmembrane protein (Pr = 0.28 | P = 6.2e-04) Gain of Sulfation at Y667 (Pr = 0.04 | P = 0.01)	ELME000079 ELME000081 ELME000084 ELME000239

Table 4 MutPred probability values of deleterious nsSNPs identified in TSHR gene.

Sr.	Mutation	P-values	Sr.	Mutation	P-values	Sr.	Mutation	P-values
1	P68S	0.672	13	L467P	0.953	25	D619G	0.873
2	T136A	0.58	14	T477P	0.861	26	L629F	0.934
3	I167N	0.871	15	T477I	0.797	27	F631L	0.931
4	N187K	0.871	16	W488R	0.956	28	D633H	0.947
5	G245S	0.695	17	S508L	0.815	29	C636G	0.945
6	S281I	0.857	18	V509A	0.57	30	P639S	0.901
7	C390W	0.791	19	R519H	0.837	31	A644V	0.577
8	C408R	0.966	20	R534C	0.382	32	I654F	0.898
9	D410N	0.686	21	A553T	0.759	33	S657R	0.716
10	G431S	0.816	22	L557W	0.848	34	F666S	0.92
11	R450H	0.62	23	M572R	0.923	35	C672Y	0.937
12	M453T	0.918	24	C600R	0.925			

TSHR Stability Prediction:

35 nsSNPs were analyzed to investigate the change in protein stability. Each of the chosen nsSNPs was delivered separately to I-mutant 2.0 and Mupro as shown in Table 5. In both cases majority of 35 nsSNPs showed destabilized effect on protein, with few increasing its stability.

Table 5 Results of TSHR protein stability prediction by I-Mutant and Mupro.

Sr.	SNP ID	Amino Acid Change	I-Mutant 2.0 prediction	RI	DDG	Mupro prediction	Mupro score
1	rs142063461	P68S	Decrease	8	-1.79	Decrease	-1.019
2	rs1888697548	T136A	Decrease	8	-0.88	Decrease	-1.273
3	rs121908862	I167N	Decrease	8	-1.66	Decrease	-1.539
4	rs2075179	N187K	Decrease	2	-0.87	Decrease	-2.262
5	rs189506473	G245S	Decrease	5	-1.43	Decrease	-0.755
6	rs121908873	S281I	Increase	5	0.23	Decrease	-0.168
7	rs121908871	C390W	Decrease	2	-0.47	Decrease	-1.383
8	rs199702292	C408R	Decrease	5	-0.72	Decrease	-0.937
9	rs121908868	D410N	Decrease	6	-1.08	Decrease	-1.132
10	rs121908883	G431S	Decrease	2	-0.89	Decrease	-0.964
11	rs189261858	R450H	Decrease	9	-1.74	Decrease	-0.933
12	rs121908864	M453T	Decrease	9	-1.65	Decrease	-1.704
13	rs121908885	L467P	Decrease	6	-2.01	Decrease	-2.158
14	rs1595178231	T477P	Decrease	3	-0.85	Decrease	-0.892
15	rs121908881	T477I	Increase	1	0.03	Decrease	-0.057
16	rs2140111252	W488R	Decrease	9	-1.54	Decrease	-1.073
17	rs772490052	S508L	Increase	2	0.47	Increase	0.115
18	rs121908874	V509A	Decrease	9	-2.84	Decrease	-1.604
19	rs780018604	R519H	Decrease	6	-0.8	Decrease	-1.141
20	rs150602845	R534C	Decrease	6	-0.5	Decrease	-0.921

21	rs121908872	A553T	Decrease	5	-0.73	Decrease	-1.33
22	rs1237819854	L557W	Decrease	6	-0.49	Decrease	-1.262
23	rs2140112304	M572R	Decrease	5	-0.53	Decrease	-1.366
24	rs121908884	C600R	Decrease	4	-0.11	Decrease	-1.346
25	rs121908859	D619G	Decrease	7	-1.27	Decrease	-1.625
26	rs121908877	L629F	Decrease	9	-0.62	Decrease	-0.642
27	rs121908861	F631L	Decrease	2	-0.62	Decrease	-0.921
28	rs28937584	D633H	Decrease	5	-0.99	Decrease	-0.623
29	rs1891823429	C636G	Decrease	9	-1.6	Decrease	-1.932
30	rs121908880	P639S	Decrease	8	-1.4	Decrease	-1.175
31	rs2140113209	A644V	Increase	2	0.21	Decrease	-0.500
32	rs767239688	I654F	Decrease	6	-1.98	Decrease	-1.484
33	rs143773384	S657R	Increase	0	-0.85	Decrease	-0.475
34	rs2140113502	F666S	Decrease	9	-2.61	Decrease	-2.072
35	rs121908875	C672Y	Decrease	5	-0.99	Decrease	-1.361

Evolutionary Conservation of TSHR protein:

To see the possible effects of the 35 seclected nsSNPs, ConSurf was used to study the conservation profile of TSHR amino acid residues. Result obtained are provided as structural representation of TSHR protein in Supplementary Figure 4. According to ConSurf prediction T136A, N187K, D410N, G431S, R450H, R519H, and D619G were predicted to be highly conserved and exposed and predicted as functional residues. Structural residues were G245S, S281I, C390W, C408R, L467P, T477P, T477I, W488R, S508L, A553T, M572R, L629F, F631L, C636G, and P639S which were predicted as highly conserved and buried. I167N, M453T, V509A, L557W, C600R, D633H, A644V, I654F, F666S, and C672Y were predicted as buried. Conservation scores for each of the selected nsSNPs are given in Table 6. These results show that nsSNPs of which locate at highly conserved region are the most damaging to the TSHR protein function and structure.

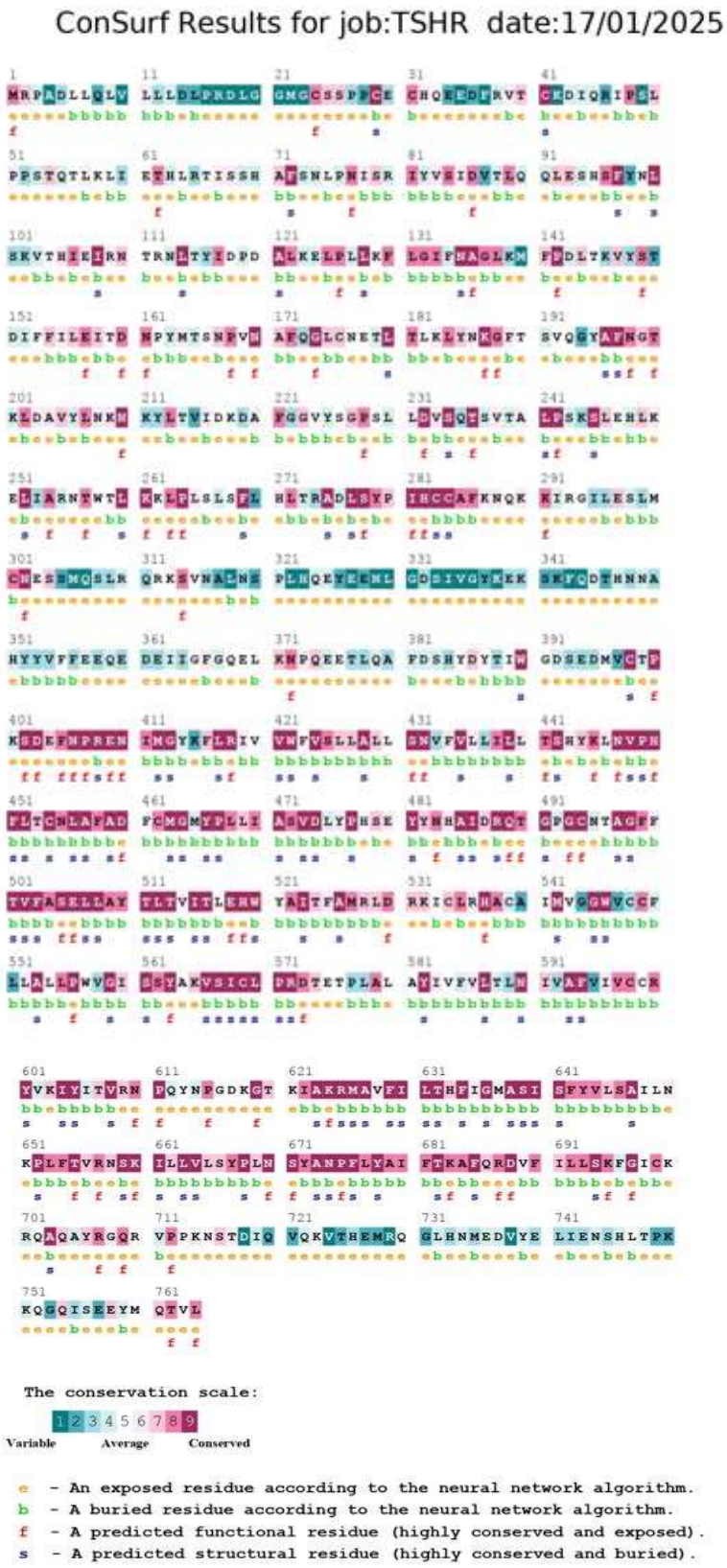


Figure 4 Represents results of ConSurf server.

Table 6 Conservation Profiles of the selected most damaging nsSNPs in TSHR.

Sr.	SNP ID	Amino Acid Change	Conservation score	Prediction
1	rs142063461	P68S	6	Exposed
2	rs1888697548	T136A	9	Highly conserved and exposed (f)
3	rs121908862	I167N	7	Buried
4	rs2075179	N187K	9	Highly conserved and exposed (f)
5	rs189506473	G245S	9	Highly conserved and buried (s)
6	rs121908873	S281I	9	Highly conserved and buried (s)
7	rs121908871	C390W	9	Highly conserved and buried (s)
8	rs199702292	C408R	9	Highly conserved and buried (s)
9	rs121908868	D410N	9	Highly conserved and exposed (f)
10	rs121908883	G431S	9	Highly conserved and exposed (f)
11	rs189261858	R450H	9	Highly conserved and exposed (f)
12	rs121908864	M453T	7	Buried
13	rs121908885	L467P	9	Highly conserved and buried (s)
14	rs1595178231	T477P	9	Highly conserved and buried (s)
15	rs121908881	T477I	9	Highly conserved and buried (s)
16	rs2140111252	W488R	9	Highly conserved and buried (s)
17	rs772490052	S508L	9	Highly conserved and buried (s)
18	rs121908874	V509A	8	Buried
19	rs780018604	R519H	9	Highly conserved and exposed (f)
20	rs150602845	R534C	7	Exposed
21	rs121908872	A553T	9	Highly conserved and buried (s)
22	rs1237819854	L557W	5	Buried
23	rs2140112304	M572R	9	Highly conserved and buried (s)
24	rs121908884	C600R	8	Buried
25	rs121908859	D619G	9	Highly conserved and exposed (f)
26	rs121908877	L629F	9	Highly conserved and buried (s)
27	rs121908861	F631L	9	Highly conserved and buried (s)
28	rs28937584	D633H	7	Buried
29	rs1891823429	C636G	9	Highly conserved and buried (s)
30	rs121908880	P639S	9	Highly conserved and buried (s)
31	rs2140113209	A644V	8	Buried
32	rs767239688	I654F	8	Buried
33	rs143773384	S657R	8	Exposed
34	rs2140113502	F666S	7	Buried
35	rs121908875	C672Y	8	Buried

3D Modelling of TSHR and Its Mutants:

First of all, phyre2 tool was used in which structure was incomplete due to amino acid missing, then Swiss-model was also used to predict structure but Ramachandran was not good for the predicted models. Then, TrRosetta was used for 3D structure generation of wild-type TSHR protein and 35 mutants. Each nsSNPs substitution was

made individually in TSHR protein sequence and then submitted to TrRosetta, which predicted their 3D structure. The wild-type structure is validated by SAVES Server v6.1 in which the ERRAT determines the quality score which was 97.24 and in RAMACHANDRAN plot 89.6% residues in the favorable region. GalaxyRefine was used for the refinement of 3D structure of TSHR wild-type, after refinement the structure was again validated by SAVES Server v6.1 in which the ERRAT was 93.58 and residues in most favored region of RAMACHANDRAN was 92.4%, suggesting great structural integrity. The result of the wild-type TSHR protein PROCHECK and ERRAT are shown in Figure 5 and 6.

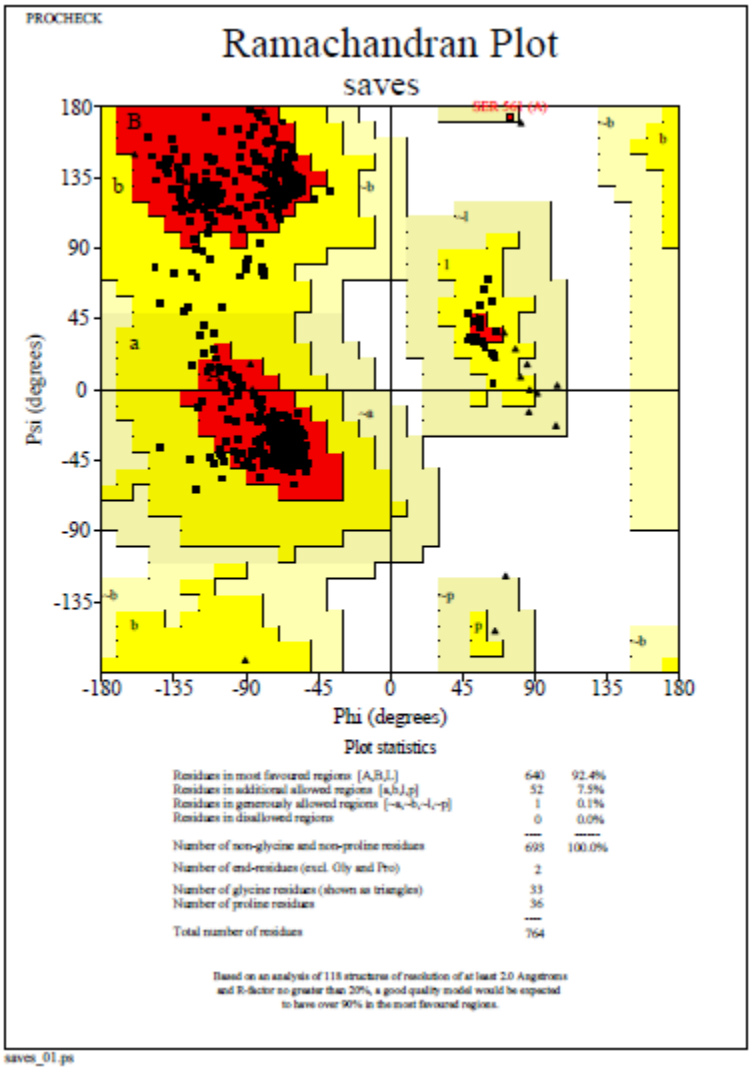


Figure 5 Represents PROCHECK Ramachandran results of wild TSHR.

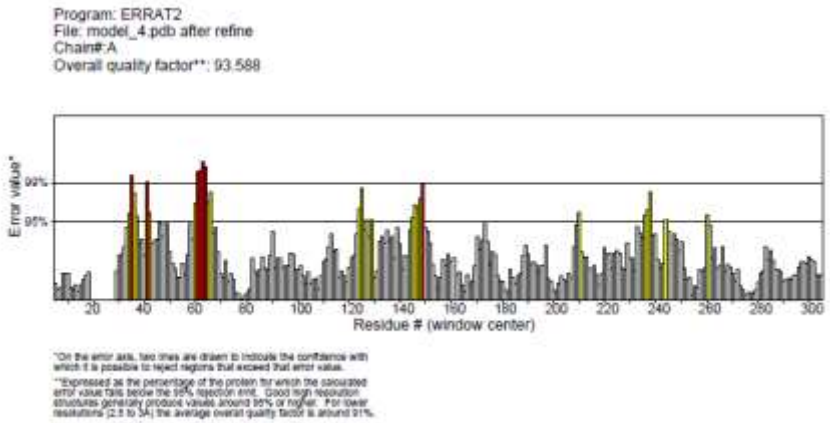


Figure 6 Represents ERRAT result of TSHR

TM-scores and RMSD were calculated for each of the mutant model. TM-score shows us the topological similarity while RMSD value show average distance between α-carbon backbones of wild and mutant models. Higher RMSD values predict greater mutant structure deviation from wild-type. Model for the mutant N187K(rs2075179) showed the greatest deviation in structure from wild-type having 3.56Å RMSD value followed by P68S (rs142063461), C390W (rs121908871), L467P (rs121908885), and W488R (rs2140111252) with 3.02Å, 2.9Å, 2,87Å, 2.81Å RMSD values, respectively. All mutants RMSD value and Tm score given in Table 7.

Table 7 RMSD TM score and values of selected nsSNPs of TSHR

SNP ID	Amino Acid Change	TM Align	RMSD
rs142063461	P68S	0.8721	3.02
rs1888697548	T136A	0.82028	1.94
rs121908862	I167N	0.8285	2.04
rs2075179	N187K	0.88892	3.56
rs189506473	G245S	0.8315	2.3
rs121908873	S281I	0.83934	2.33
rs121908871	C390W	0.88822	2.9
rs199702292	C408R	0.81795	2.09
rs121908868	D410N	0.82025	1.95
rs121908883	G431S	0.81412	1.84
rs189261858	R450H	0.83258	2.66
rs121908864	M453T	0.85034	2.65
rs121908885	L467P	0.88124	2.87
rs1595178231	T477P	0.86836	2.64
rs121908881	T477I	0.8216	2.14
rs2140111252	W488R	0.84318	2.81
rs772490052	S508L	0.86823	2.38

rs121908874	V509A	0.82719	2.23
rs780018604	R519H	0.82385	1.98
rs150602845	R534C	0.83678	2.5
rs121908872	A553T	0.81097	1.68
rs1237819854	L557W	0.89433	2.44
rs2140112304	M572R	0.82244	2.02
rs121908884	C600R	0.9009	2.07
rs121908859	D619G	0.83956	2.23
rs121908877	L629F	0.83492	2.13
rs121908861	F631L	0.83211	2.54
rs28937584	D633H	0.81084	2.24
rs1891823429	C636G	0.83632	2.74
rs121908880	P639S	0.83477	2.78
rs2140113209	A644V	0.90506	2.06
rs767239688	I654F	0.82286	2.16
rs143773384	S657R	0.8338	1.89
rs2140113502	F666S	0.82466	2.34
rs121908875	C672Y	0.83836	1.96

The mutants having greater RMSD values (N187K, P68S, C390W, L467P, and W488R) were selected and superimposed over native protein for further analysis using PyMol shown in Fig 7. All the mutants and Wild-Type SAVES Server v6.1 result shown in Table 8.

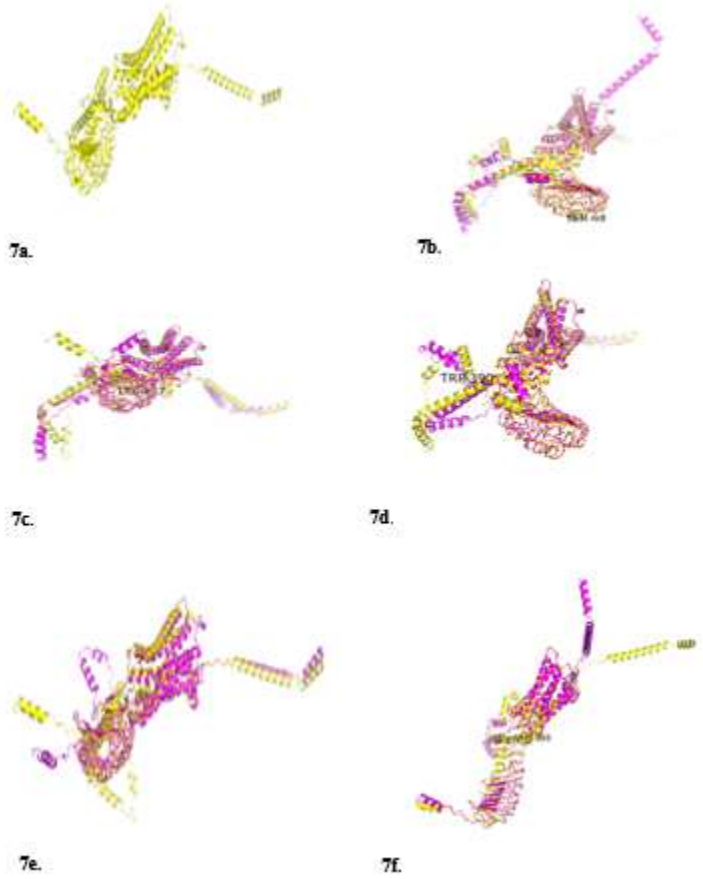


Figure 7

(7a) Structure of wild type (yellow) TSHR protein. (7b) Superimposed wild type TSHR protein and its mutant having Proline(P) to Serine(S) at position 68. (7c) Superimposed wild type TSHR protein and its mutant having Asparagine(N) to Lysine(K) at position at position 187. (7d) Superimposed wild TSHR type protein and its mutant Cysteine) to Tryptophan(W) at position 390. (7e) Superimposed wild type TSHR protein and its mutant having Leucine(L) to Proline(P) at position 467. (7f) Superimposed wild type TSHR protein and its mutant having Tryptophan(W) to Arginine(R) at position 488

Table 8 SAVES server results of wild and Mutants.

SNPs	Amino Acid Change	Errat Score	Ramachandran Plot
	TSHR Before Refinement	97.24	89.60%
	TSHR After Refinement	93.58	92.40%
rs142063461	P68S	96.7568	90.20%
rs2075179	N187K	99.0541	90.00%
rs121908871	C390W	98.489	89.30%
rs121908885	L467P	99.0358	88.90%
rs2140111252	W488R	97.6808	90.30%

Predicted Post Translational Modification sites (PTMs):

PTMs have a crucial role in controlling the structure and functions of protein. In this study, we sought to investigate whether the high risk nsSNPs have any effect on PTMs in TSHR.

Methylation:

The predicted methylation sites and their respective score are presented in Table 9. GPS-MSP 3 identified a total of 10 methylation sites, with position 710 showing the highest methylation score of 20.58, the highest among all predicted sites. PRmePRed predicted 6 methylation sites within TSHR protein, with position 707 displaying the highest score 0.85.

Table 9 TSHR Methylation results by PRmePRed and GPS-MSP 3.0

PRmePRed			GPS-MSP 3.0			
R site	Peptides	Prediction Score	Position	Code	Type	Score
17	QLVLLLDLPRDLGGMGCSS	0.832958	2	R	R.all	3.5
293	CAFKNQKKIRGILESLMCN	0.577832	80	R	R.all	5.24
310	CNESSMQSLRQRKSVNALN	0.753781	211	K	K.all	4.9
312	ESSMQSLRQRKSVNALNSP	0.622816	261	K	K.all	1.87
707	ICKRQAQAYRGQRVPPKNS	0.855149	274	R	R.all	3.95
710	RQAQAYRGQRVPPKNSTDI	0.555137	287	K	K.all	8.69
			313	K	K.all	10.65
			519	R	R.all	2.86
			603	K	K.all	1.37
			710	R	R.all	20.58

Phosphorylation

Phosphorylation sites in the TSHR protein were predicted by GPS 6.0 and NetPhos 3.1. The predicted phosphorylation sites and their respective details are summarized in Table 10. NetPhos 3.1 identified a total of 74 phosphorylation sites, while GPS 6.0 predicted 16 phosphorylation sites, consisting 4 serine sites and 12 tyrosine sites.

Table 10 TSHR Phosphorylation results by NetPhos 3.1 and GPS 6.0

Residues	NetPhos 3.1			GPS 3.0			
	Position	Score	Kinase	Position	Score	Cutoff	kinase
Serine (S)	25	0.554	cdc2	308	0.1004	0.0996	CK1
	49	0.848	unsp	314	0.0413	0.04	CAMK
	49	0.562	PKA	314	0.3324	0.0996	CK1
	49	0.545	DNAPK	393	0.1712	0.0996	CK1
	53	0.503	GSK3	716	0.1562	0.0996	CK1
	94	0.515	cdc2				
	96	0.623	unsp				
	101	0.529	cdc2				
	166	0.506	cdc2				
	229	0.592	unsp				
	234	0.627	DNAPK				
	234	0.517	cdc2				

	234	0.513	ATM				
	237	0.968	unsp				
	266	0.754	unsp				
	266	0.739	PKA				
	268	0.539	cdc2				
	278	0.931	unsp				
	278	0.75	PKC				
	281	0.856	unsp				
	298	0.6	PKC				
	308	0.914	PKC				
	308	0.65	unsp				
	314	0.635	PKA				
	314	0.567	PKC				
	314	0.536	RSK				
	320	0.515	p38MAPK				
	333	0.612	unsp				
	333	0.54	cdc2				
	333	0.537	CKI				
	341	0.77	unsp				
	383	0.94	unsp				
	393	0.545	CKII				
	402	0.952	unsp				
	425	0.705	PKA				
	472	0.517	PKA				
	472	0.501	cdc2				
	508	0.538	PKA				
	562	0.769	PKC				
	562	0.582	DNAPK				
	646	0.652	PKC				
	659	0.513	PKA				
	694	0.573	PKC				
	716	0.519	PKA				
	745	0.503	CKI				
	756	0.893	unsp				
Threonine (T)	40	0.923	unsp				
	40	0.723	PKC				
	54	0.593	DNAPK				
	56	0.825	unsp				
	56	0.772	PKC				
	66	0.514	cdc2				
	104	0.548	PKG				
	111	0.577	unsp				
	111	0.568	PKA				
	145	0.565	PKC				

	150	0.56	PKC				
	159	0.544	cdc2				
	159	0.507	CKII				
	165	0.744	unsp				
	181	0.728	unsp				
	181	0.709	PKC				
	214	0.799	PKC				
	236	0.58	unsp				
	236	0.507	cdc2				
	239	0.517	cdc2				
	239	0.501	PKC				
	257	0.813	PLC				
	257	0.657	unsp				
	259	0.89	PKC				
	259	0.697	unsp				
	388	0.559	unsp				
	399	0.983	unsp				
	399	0.561	p38MAPK				
	441	0.74	PKC				
	477	0.553	CKII				
	490	0.625	PKC				
	524	0.585	PKC				
	574	0.501	unsp				
	607	0.834	PKC				
	620	0.665	PKC				
	620	0.548	PKG				
	655	0.7	PKC				
	659	0.513	PKA				
	682	0.643	PKC				
	748	0.97	unsp				
	748	0.65	PKC				
	748	0.56	p38MAPK				
	748	0.503	cdk5				
Tyrosine (Y)	82	0.659	unsp	116	0.9636	0.897	TK
	116	0.97	unsp	163	0.96	0.897	TK
	116	0.511	INSR	225	0.9656	0.897	TK
	163	0.577	INSR	337	0.9094	0.897	TK
	195	0.873	unsp	352	0.9925	0.897	TK
	212	0.797	unsp	353	0.9918	0.897	TK
	279	0.629	unsp	476	0.9567	0.897	TK
	385	0.521	unsp	482	0.9229	0.897	TK
	387	0.764	unsp	582	0.9337	0.897	TK
	476	0.788	unsp	613	0.9463	0.897	TK
	476	0.527	INSR	739	0.9993	0.897	TK

	481	0.715	unsp	759	0.974	0.897	TK
	481	0.504	INSR				
	613	0.67	unsp				
	759	0.687	unsp				

Glycosylation

N-Glyde predicted 7 sites to be N-glycosylated, while NetOglyc 4.0 predicted 101 sites for glycosylation as represented in Table 11.

Table 11 TSHR Glycosylation results by N-Glyde and NetOglyc 4.0

N-Glyde		NetOglyc4.0	
Position	Score	Position	Score
77	0.666	25	0.231718
99	0.7387	26	0.266689
113	0.6722	40	0.063039
177	0.7619	49	0.529642
198	0.6812	53	0.423991
302	0.5573	54	0.10359
715	0.5137	56	0.468991
		62	0.284991
		66	0.185954
		69	0.34561
		73	0.454385
		79	0.045401
		84	0.197671
		88	0.09481
		94	0.209873
		96	0.23119
		101	0.098691
		104	0.2672
		111	0.621103
		115	0.24298
		136	0.218451
		145	0.029114
		149	0.106357
		150	0.097444
		159	0.305406
		165	0.161553
		166	0.17809
		179	0.022805
		181	0.027744
		190	0.055776
		191	0.070527
		200	0.034991
		214	0.161277

		226	0.08534
		229	0.548637
		234	0.180145
		236	0.439864
		237	0.280995
		239	0.420144
		243	0.414679
		257	0.150673
		259	0.287266
		266	0.37429
		268	0.121961
		273	0.401611
		278	0.068655
		281	0.160216
		298	0.15339
		304	0.31008
		305	0.418828
		308	0.426776
		314	0.749188
		320	0.25532
		333	0.271943
		341	0.339299
		346	0.238966
		377	0.367013
		383	0.444376
		388	0.081006
		393	0.167343
		399	0.239567
		402	0.132237
		425	0.015085
		441	0.008158
		442	1.58E+00
		472	0.012288
		477	0.025203
		479	0.027379
		490	0.021053
		496	0.018312
		501	0.021872
		505	0.027617
		508	0.005707
		511	0.012988
		513	1.66E+00
		516	0.005679
		524	0.008041
		561	0.083934

		562	0.073804
		567	0.093145
		574	0.059508
		576	0.020164
		588	0.014874
		607	0.010174
		620	0.009405
		632	0.005233
		641	0.00449
		646	0.009545
		655	0.024909
		657	0.129018
		659	0.015018
		671	0.003875
		682	4.22E+00
		694	0.007584
		716	0.463055
		717	0.342546
		725	0.304911
		745	0.250252
		748	0.122283
		756	0.06145
		762	0.00651

Ubiquitylation

GPS-Uber predict 3 ubiquitylation sites at position 42, 250 and 714. Site 42 reflected a high probability of modification with a score of 0.5608. Predicted site results along with their peptide sequence and scores represented in Table 12.

Table 12 TSHR Ubiquitylation by GPS-Uber

GPS-Uber					
Position	**Code**	**Kinase**	**Peptide**	**Score**	**Cutoff**
42	k	General	HQEEDFRVTCKDIQRIPSLPP	0.5608	0.3566
250	K	General	ALPSKGLEHLKELIARNTWTL	0.426	0.3566
714	K	General	QAYRGQRVPPKNSTDIQVQKV	0.4926	0.3566

Gene-gene interaction of TSHR

GeneMANIA results showed that the physical interaction of TSHR is with ASPHD1, TMEM19, SCRIB, LAMA3, SLC26A7, TMCO3, LHCGR, PTGDR, TG, GNAS, CGA, CALCR, PTH, TSHB, FSHR, GNA12, HSPA5, CALR. TSHR is co-expressed with SLC26A4, TMEM19, SLC26A7, TPO, PTGDR, LHCGR, TG, PTH, FSHR, GNA12, CALR. It is co-localized with TPO, SLC26A4, SLC26A7, TG, and PTH. In pathway, it has a relation with TSHB, TG, CGA, and GNAS. Genetic interactions are with FSHR and LHCGR and share protein domain with FSHR, LHCGR, and PTGDR. STRING predicts cumulative score that are shown in Table 13. Gene interactions are expected by GeneMANIA and STRING are shown in Figure 8 and 9 respectively.

Table 13 STRING predicts a cumulative score for each gene with TSHR

Interacting Gene	Combined Score	Evidence Suggesting a functional Links
TSHB	0.998	Expression: Yes 0.900, Databases: Yes 0.900
GNAS	0.997	Expression: Yes 0.900, Databases: Yes 0.900
TG	0.972	None/ Insignificant
TPO	0.971	None/ Insignificant
GNAQ	0.964	Databases: Yes 0.900
CGA	0.964	Expression: Yes 0.900, Databases: Yes 0.400
NKX2-1	0.954	None/ Insignificant
IGF1R	0.951	None/ Insignificant
FSHB	0.95	Expression: None, Databases: yes 0.900
SLC5A5	0.947	None/ Insignificant

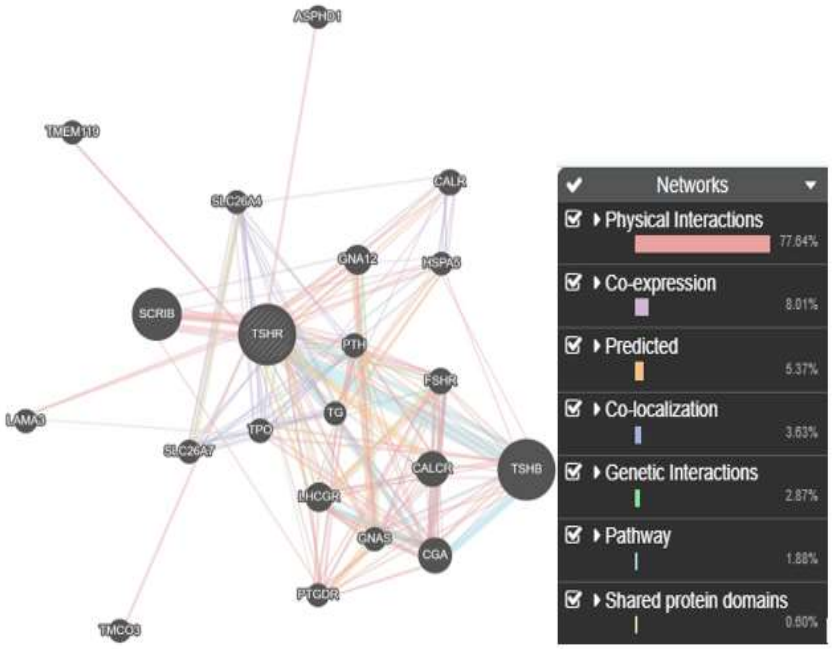


Figure 8 Gene-gene interaction of TSHR with other genes proposed by GeneMANIA

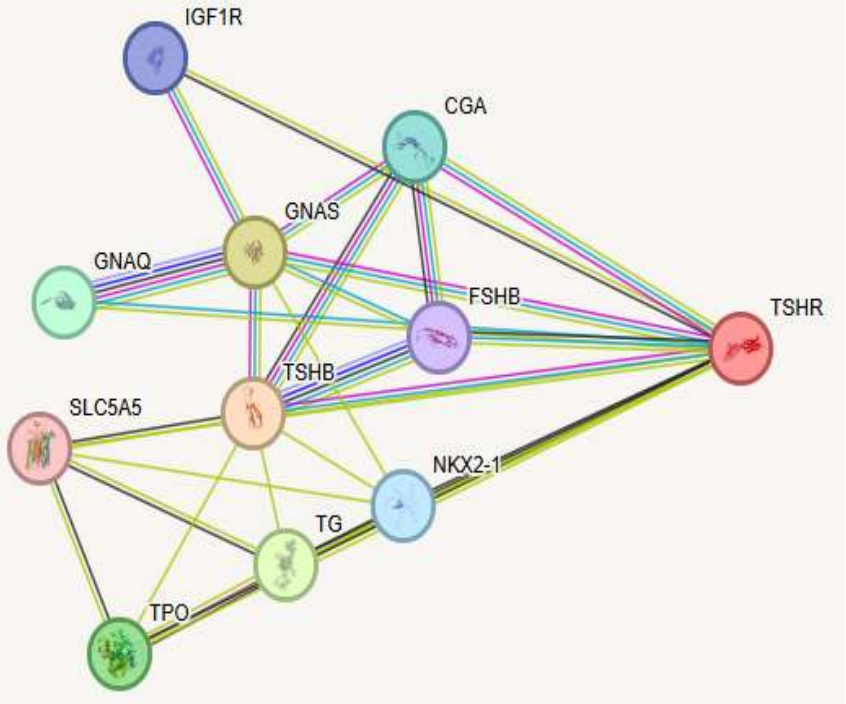


Figure 9 Gene-gene interaction of TSHR with other genes predicted by STRING

Discussion

Identifying biological relevant SNPs can help in developing SNP-based genetic profile that can be used as genetic screening markers to determine an individual's risk of acquiring various diseases and to investigate inheritance patterns (45). Genetic variations in TSHR are directly involved in the pathogenesis of autoimmune hypothyroidism, Graves' disease, toxic adenomas, familial and sporadic non-autoimmune hypothyroidism, and forms of resistance to TSH. TSHR is a vital thyrocyte membrane protein in the thyroid gland and a risk factor for TSH abnormalities (46). Genetic polymorphism within the TSHR region may influence TSHR structure, expression and/or post-translational processing.

The current research work identified the impact of nsSNPs of the TSHR gene on the structural and functional aspects using various in silico tools. This analysis commenced with the separation of 81 nsSNPs from the 80237 SNPs of TSHR gene collected from the dbSNP database of NCBI. Then, the TSHR protein sequence, accumulated from the UniProt database for conducting further analysis to identify the deleterious nsSNPs in TSHR gene. The functional and structural changes in protein occurred due to highly damaging nsSNPs, which were from the results acquired from various servers. In SIFT, 47 deleterious and 34 neutral mutations were found among the total of 81 missense SNPs. While in the PROVEAV, 42 and 39 mutations showed a deleterious and neutral effect on protein, respectively. On the other hand, PhD-SNP server showed 45 deleterious and 36 neutral mutations. The SNP&GO server displayed 62 nsSNPs were disease causing among the total of 81 and then, all the deleterious nsSNPs from the above-mentioned tools were carried out to Polyphen2 server which showed that 32 nsSNPs were probably damaging, 1 benign and 2 were possibly damaging among 35 amino acid substitutions. Then MutPred was used which predicted the statistical significance of nsSNPs on the structure and function of TSHR protein, where differentiated characteristics of deleterious nsSNPs were revealed (47). Afterward, structural stability changes of TSHR protein were analyzed because the protein stability has a significant effect on the function and activity of proteins(48). Here, I-Mutant 2.0 and MUpro both were used for the prediction of stability changes of

TSHR. In the I-Mutant 2.0 server, 30 mutations with decreased stability and 5 mutations with increased stability were found. While in the Mupro server, 34 and 1 mutations with decreased and increased stability, respectively, were found.

Then the results obtained from the above eight servers were analyzed and compared, which eventually assorted 35 potentially deleterious mutations for TSHR protein. These 35 deleterious mutations went through the ConSurf server, which determined the conservational score of the TSHR protein for different amino acid residues, where 22 residues with score 9, 6 residues with score 8, 5 residues with score 7, 1 residue with score 6 and 1 residues with score 5. In conservational analysis, score 1 denotes the least conserved or variable and score 9 denotes the highly conserved amino acid (49).

The above analysis thus concludes that these nsSNPs might cause maximum damage to the TSHR protein by affecting its stability. However, the 3D structure of the protein plays an important role in understanding the overall effects of nsSNPs on protein function. In this study, we assessed the structural impact of 35 damaging nsSNPs in the TSHR protein using TrRosetta to generate 3D models for both the wild-type and mutant proteins. Structural validation of the wild-type TSHR protein was performed using the SAVES Server v6.1, yielding an ERRAT score of 97.24, indicating high structural quality, and 89.6% of the residues were positioned in the favorable region of the Ramachandran plot. Further refinement of the wild-type structure with GalaxyRefine led to an ERRAT score of 93.58 and 92.4% of residues in the most favored Ramachandran region, suggesting robust structural integrity.

To evaluate the effects of the nsSNPs, we calculated TM-scores and RMSD values for each mutant model. The N187K (rs2075179) mutant showed the highest deviation (3.56 Å RMSD), followed by P68S (rs142063461), C390W (rs121908871), L467P (rs121908885), and W488R (rs2140111252), with RMSD values of 3.02 Å, 2.9 Å, 2.87 Å and 2.81 Å, respectively. These mutants, which displayed significant structural deviations, were selected for further analysis and superimposed with the native protein using PyMol software.

In this study, various post-translational modifications (PTMs) in the TSHR protein were predicted using specialized tools. Methylation sites were identified by GPS-MSP and PRmePRed, with GPS-MSP predicting 10 sites, the highest score of 20.58 being at position 710. PRmePRed predicted 6 sites, with position 707 showing the highest score of 0.85. Phosphorylation sites were predicted by GPS 6.0 and NetPhos 3.1, with NetPhos 3.1 identifying 74 sites and GPS 6.0 predicting 16 sites, including 4 serine and 12 tyrosine residues. Glycosylation sites were predicted by N-Glyde and NetOglyc 4.0, with N-Glyde identifying 7 N-glycosylation sites and NetOglyc 4.0 predicting 101 glycosylation sites. Ubiquitylation prediction using the GPS-Uber server identified 3 lysine modification sites at positions 42, 250, and 714, with position 42 showing the highest modification probability (score 0.5608).

GeneMANIA and STRING analysis revealed key interactions of TSHR with several genes, including ASPHD1, TMEM19, LHCGR, FSHR, TG, GNAS AND PTH. TRSHR was found to be expressed and co-localized with genes like SLC26A4, TPO and TG and it shares genetic interactions with FSHR and LHCGR. Pathway analysis highlighted associations with TSHB, TG, CGA, and GNAS. Thus, it can be concluded from their interaction patterns and co-expression profiles that any of the most damaging nsSNPs in TSHR gene will ultimately effect and disturb normal functioning of other interacting genes. STRING further confirmed these interactions with a cumulative score, suggesting TSHR involvement in a complex network related to thyroid hormone signaling. Optimum protein-protein interaction is important for maintenance of normal cell function and to allow for metabolite interaction(50).

As our study was in depth, it provides all the analysis and information that are required for the investigation of most deleterious nsSNPs. Like ours, there are certain limitations in every research. This study is based on computational servers which are based on statistical and mathematical algorithms. Therefore, to confirm

these results, laboratory investigation is required. Our study provides an insight about nsSNPs of TSHR gene, its protein 3D structure, its PTM potential sites and its gene-gene interaction with other gene which might be helpful in further studies of TSHR in order to better understanding its role in thyroid related genetic disorders.

Conclusion:

This study identified 5 major high risk nsSNPs, rs142063461 (P68S), rs2075179 (N187), rs121908871 (C390W), rs121908885 (L467), rs2140111252 (w488R) within the coding region of TSHR gene. They may have role in thyroid disorders associated with TSHR gene as they are involved in decreasing the protein stability and have higher RMSD values. The mutants possessing these nsSNPs showed deviation in structure from wild type TSHR protein. These nsSNPs can be significance for the therapeutic strategies and personalized medicine and can be used for further experimental investigations to study the role of these nsSNPs in pathogenesis of related diseases.

Author Contribution statement

JN write and prepared the original draft. SF, SR, AF and MU proofread and edited the work. FI supervised the work. All the authors read and approved the final manuscript.

Ethical Approval

Not applicable.

Competing interests

None.

References

1. Vasileiou M, Gilbert J, Fishburn S, Boelaert K. Thyroid disease assessment and management: summary of NICE guidance. Bmj. 2020;368.
2. Zhang X, Wang X, Hu H, Qu H, Xu Y, Li Q. Prevalence and trends of thyroid disease among adults, 1999-2018. Endocrine Practice. 2023;29(11):875-80.
3. Mense MG, Boorman GA. Thyroid gland. Boorman's pathology of the rat: Elsevier; 2018. p. 669-86.
4. Pirahanchi Y, Tariq MA, Jialal I. Physiology, thyroid. StatPearls [Internet]. 2023.
5. Makkonen K, Jännäri M, Crisóstomo L, Kuusi M, Patyra K, Melnyk V, et al. Mechanisms of thyrotropin receptor–mediated phenotype variability deciphered by gene mutations and M453T-knockin model. JCI insight. 2024;9(4):e167092.
6. Pujol-Borrell R, Giménez-Barcons M, Marín-Sánchez A, Colobran R. Genetics of Graves' disease: special focus on the role of TSHR gene. Hormone and Metabolic Research. 2015;47(10):753-66.
7. Zhang Y, Tan Y, Zhang Z, Cheng X, Duan J, Li Y. Targeting Thyroid-Stimulating Hormone Receptor: A Perspective on Small-Molecule Modulators and Their Therapeutic Potential. Journal of Medicinal Chemistry. 2024;67(18):16018-34.
8. Begum MN, Mahtarin R, Islam MT, Ahmed S, Konika TK, Mannoor K, et al. Molecular investigation of TSHR gene in Bangladeshi congenital hypothyroid patients. PloS one. 2023;18(8):e0282553.
9. Akhtar M, Jamal T, Jamal H, Din JU, Jamal M, Arif M, et al. Identification of most damaging nsSNPs in human CCR6 gene: In silico analyses. International journal of immunogenetics. 2019;46(6):459-71.
10. Nadeali Z, Mohammadi-Zaniani Z, Biglari S, Molavi N, Zardoui K, Mirfendereski S, et al. Investigating TSHR gene variants in consanguineous families: novel insights into variable expression in familial congenital hypothyroidism. Frontiers in Endocrinology. 2025;16:1559281.
11. Zhang H-Y, Wu F-Y, Li X-S, Tu P-H, Zhang C-X, Yang R-M, et al. TSHR Variant Screening and Phenotype Analysis in 367 Chinese Patients With Congenital Hypothyroidism. Annals of Laboratory Medicine. 2024;44(4):343-53.

12. Stoupa A, Carré A, Polak M, Szinnai G, Schoenmakers N. Genetics of primary congenital hypothyroidism: three decades of discoveries and persisting etiological challenges. European Thyroid Journal. 2025;14(2).
13. Sabiha B, Bhatti A, Roomi S, John P, Ali J. In silico analysis of non-synonymous missense SNPs (nsSNPs) in CPE, GNAS genes and experimental validation in type II diabetes mellitus through Next Generation Sequencing. Genomics. 2021;113(4):2426-40.
14. Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, et al. Ensembl variation resources. Database. 2018;2018:bay119.
15. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols. 2009;4(7):1073-81.
16. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nature protocols. 2016;11(1):1-9.
17. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic acids research. 2012;40(W1):W452-W7.
18. Fialho LF, de Alencar SA. Use of computational tools to assess the functional impact of missense variants in the CYP2D6 gene. Brazilian Journal of Health Review. 2024;7(5):e73137-e.
19. Kamal MM, Teeya ST, Rahman MM, Talukder MEK, Sarmin S, Wani TA, et al. Prediction and assessment of deleterious and disease causing nonsynonymous single nucleotide polymorphisms (nsSNPs) in human FOXP4 gene: An in silico study. Heliyon. 2024.
20. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 2006;22(22):2729-34.
21. Abdalla MH, Babeker MM. IN SILICO ANALYSIS OF CODING NON-SYNONYMOUS SINGLE NUCLEOTIDE POLYMORPHISM (SNP) OF HUMAN NQO1 GENE AND THEIR IMPACT ON BENZENE INDUCED HAEMATOTOXICITY. 2022.
22. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. BMC genomics. 2013;14:1-7.
23. Ahmed EM, Elangeeb ME, Adam KM, Abuagla HA, MohamedAhmed AAE, Ali EW, et al. Computational Analysis of Deleterious nsSNPs in INS Gene Associated with Permanent Neonatal Diabetes Mellitus. Journal of Personalized Medicine. 2024;14(4):425.
24. Waheed S, Ramzan K, Ahmad S, Khan MS, Wajid M, Ullah H, et al. Identification and In-Silico study of non-synonymous functional SNPs in the human SCN9A gene. Plos one. 2024;19(2):e0297367.
25. Manfredi M, Savojardo C, Martelli PL, Casadio R. E-SNPs&GO: embedding of protein sequence and function improves the annotation of human pathogenic variants. Bioinformatics. 2022;38(23):5168-74.
26. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature methods. 2010;7(4):248-9.
27. Itan Y, Shang L, Boisson B, Ciancanelli MJ, Markle JG, Martinez-Barricarte R, et al. The mutation significance cutoff: gene-level thresholds for variant predictions. Nature methods. 2016;13(2):109-10.
28. Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic acids research. 2005;33(suppl_2):W306-W10.

29. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. Proteins: Structure, Function, and Bioinformatics. 2011;79(3):830-8.
30. Tariq H, Amir A, Saleem M, Ramzan K, Aslam T, Asif M. In-Depth In-Silico Functional, And Structural Screening Of IL-4 Gene Variants Linked with Pink Eye Infection. Int J Sci R Tech. 2024;1(11).
31. Arifuzzaman M, Mitra S, Das R, Hamza A, Absar N, Dash R. In silico analysis of nonsynonymous single-nucleotide polymorphisms (nsSNPs) of the SMPX gene. Annals of human genetics. 2020;84(1):54-71.
32. Kumar H, Kim P. Artificial intelligence in fusion protein three-dimensional structure prediction: Review and perspective. Clinical and Translational Medicine. 2024;14(8):e1789.
33. Park T, Won J, Baek M, Seok C. GalaxyHeteromer: protein heterodimer structure prediction by template-based and ab initio docking. Nucleic Acids Research. 2021;49(W1):W237-W41.
34. Agnihotry S, Pathak RK, Singh DB, Tiwari A, Hussain I. Protein structure prediction. Bioinformatics: Elsevier; 2022. p. 177-88.
35. Bharadwaj KK, Ahmad I, Pati S, Ghosh A, Sarkar T, Rabha B, et al. Potent bioactive compounds from seaweed waste to combat cancer through bioinformatics investigation. Frontiers in Nutrition. 2022;9:889276.
36. Deng W, Wang Y, Ma L, Zhang Y, Ullah S, Xue Y. Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins. Briefings in Bioinformatics. 2017;18(4):647-58.
37. Kumar P, Joy J, Pandey A, Gupta D. PRmePRed: A protein arginine methylation prediction tool. PLoS One. 2017;12(8):e0183318.
38. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of molecular biology. 1999;294(5):1351-62.
39. Iqbal MW, Shahab M, Zheng G, Sun X, Yuan Q, Almaary KS, et al. Analysis of damaging non-synonymous SNPs in GPx1 gene associated with the progression of diverse cancers through a comprehensive in silico approach. Scientific Reports. 2024;14(1):28690.
40. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KTB, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal. 2013;32(10):1478-88.
41. Joshi HJ, Gupta R. Eukaryotic glycosylation: online methods for site prediction on protein sequences. Glycoinformatics. 2015:127-37.
42. Sivakumar K, Subbiah U. Computational analysis of non-synonymous SNPs in the human LCN2 gene. Egyptian Journal of Medical Human Genetics. 2024;25(1):94.
43. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic acids research. 2010;38(suppl_2):W214-W20.
44. Shaukat Z, Aiman S, Li C-H. Protein-protein interactions: Methods, databases, and applications in virus-host study. World Journal of Virology. 2021;10(6):288.
45. Subbiah HV, Babu PR, Subbiah U. Determination of deleterious single-nucleotide polymorphisms of human LYZ C gene: an in silico study. Journal of Genetic Engineering and Biotechnology. 2022;20(1):92.
46. Pang Y, Guan Y, Jin X, Shen H, Liu L, Jia Q, et al. Association of TSHR gene copy number variation with TSH abnormalities. Biological Trace Element Research. 2018;186:85-90.

47. Laskar FS, Bappy MNI, Hossain MS, Alam Z, Afrin D, Saha S, et al. An In silico Approach towards Finding the Cancer-Causing Mutations in Human MET Gene. International Journal of Genomics. 2023;2023(1):9705159.
48. Marabotti A, Scafuri B, Facchiano A. Predicting the stability of mutant proteins by computational approaches: an overview. Briefings in Bioinformatics. 2021;22(3):bbaa074.
49. Kamal MM, Shantanu KFH, Teeya ST, Rahman MM, Hasan AM, Chivers DP, et al. Investigating the functional and structural effect of non-synonymous single nucleotide polymorphisms in the cytotoxic T-lymphocyte antigen-4 gene: An in-silico study. PloS one. 2025;20(1):e0316465.
50. Xu Y, Wang H, Nussinov R, Ma B. Protein charge and mass contribute to the spatio-temporal dynamics of protein–protein interactions in a minimal proteome. Proteomics. 2013;13(8):1339-51.

RJNMSR
Review Journal of Neurological & Medical Sciences Review