Bulletin of Biological and Allied Sciences Research
***ISSN:* 2521-0092**
www.bbasr.org
DOI: https://doi.org/10.64013/bbasr.v2026i1.132
*Bull. Biol. All. Sci. Res., Volume, **11**: 132*
Original Research Article ***Open Access***

INTEGRATED *IN SILICO* AND MOLECULAR DOCKING ANALYSIS OF GDF7 nsSNPs ASSOCIATED WITH BARRETT'S ESOPHAGUS

BIBI R[1], RIAZ S[2], KHAN S[3], SALMANOVA R[4], SADIKAN MZ[5], MOEEZ S[6*], KIRBAG S[7], ALI Q[8], ULLAH A[9*]

[1]*Department of Bioinformatics, Faculty of Computing and Information Technology (FCIT), International Islamic University (IIUI), H-10, Islamabad, Pakistan*
[2]*Scientific Adviser, Ministry of Science and Technology, Islamabad, Pakistan*
[3]*Department of Microbiology, Hazara University Mansehra, Pakistan*
[4]*Faculty of Natural Sciences and Agriculture, Department of Biology, Nakhchivan State University, Nakhchivan, Azerbaijan*
[5]*Faculty of Pharmacy and Health Sciences, Universiti Kuala Lumpur Royal College of Medicine Perak, Jalan Greentown, 30450 Ipoh, Perak, Malaysia*
[6]*Department of Biological Sciences, International Islamic University (IIUI), H-10, Islamabad, Pakistan.*
[7]*Department of Biology, Faculty of Science, Firat Univerity Firat 23119 Elazig, Turkey*
[8]*Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX. 54590, Lahore, Pakistan*
[9]*Department of Allied Health Sciences, Iqra National University, Peshawar 25000, Pakistan*
**Correspondence Author Email Address: sadaf.haris2399@gmail.com; aminbiotech7@gmail.com*

(Received, 4th February 2026, Accepted 12th July 2026, Published 25th July 2026)

Abstract *Growth differentiation factor 7 (GDF7), a member of the transforming growth factor-β superfamily, plays a vital role in cell differentiation, survival, and development. Mutations in GDF7 have been linked to various disorders, including Barrett's esophagus (BE), a condition arising from chronic gastroesophageal reflux disease (GERD). This study aimed to evaluate the impact of non-synonymous single nucleotide polymorphisms (nsSNPs) on the structure, stability, and function of the GDF7 protein and the interaction with other protein followed by identifying selective inhibitors to modulate the relevant signaling pathway associated with the identified harmful variant. Missense SNPs were analyzed using SIFT, MUpro, PolyPhen-2, PhD-SNP, GeneMANIA, Meta-SNP, I-mutant, Consurf, Project Hope and SNPs&GO. The wild-type structure of GDF7 was predicted using the AlphaFold database, while mutant models were generated using Swiss Model. Binding sites of GDF7 were predicted using COACH-D, and molecular docking anslalysis was conducted using PyRx. A total of 325 missense SNPs were found in the GDF7 gene, of which five nsSNPs (rs146317060, rs146567848, rs367794645, rs369544272, rs377143734) were specified as damaging to the structure, stabilization, and role of the GDF7 protein and were housed in conserved regions. The selected 5 mutations (D364H, R419H, K352T, V434A, I423M) were incorporated in the modeling of the mutant structure and template Q9BDW8.1.A model from SWISS-MODEL was selected as most suitable, with 94.63% identity with the amino acid sequence provided. Docking score and interaction analysis showed that Proflavine has strong binding affinity and a variety of interactions with mutated GDF7, so this suggests that the compound has promising potential and can be further developed as a drug for targeting diseases caused by mutations in the GDF7 pathway, when coupled with experimental results in vitro. This study identified 5 harmful nsSNPs in GDF7 and revealed Proflavine as a potential therapeutic molecule for targeting Barrett's Esophagus. The findings provide insights into personalized medicine approaches for managing diseases influenced by GDF7 alterations*.

[*Citation:* Bibi, R., Riaz, S., Khan, S., Salmanova, R., Sadikan, M.Z., Moeez, S., Kirbag, S., Ali, Q., Ullah, A. (2026). Integrated *in silico* and molecular docking analysis of GDF7 nsSNPs associated with Barrett's esophagus. *Bull. Biol. All. Sci. Res.* **11***: 132.* https://doi.org/10.64013/bbasr.v2026i1.132]
Keywords*:* *Growth differentiation factor 7; Non-synonymous SNPs; Barrett's esophagus; Molecular Docking; Protein Stability; Proflavine*

Introduction

Growth differentiation factor 7 (GDF7) is a secreted signaling molecule belonging to the TGFβ (Transforming Growth Factor β) which is proteins' superfamily. SMAD (Suppressor of Mothers against Decapentaplegic) family transcription factors that control the expression of genes are recruited and activated by ligands of this TGFβ family(Yang *et al.*, 2026). GDF7 regulates diverse physiological activities like specialization and survival of cells, tumor formation, and development of the embryo (Tao *et al.*, 2023). Studies have shown that the expression of the GDF7 gene is in the brain, and it plays an important role in the motor region of the primate neocortex and roof plate (Tao *et al.*, 2023).

Through generating signals within the roof plate, GDF7 particularly functions to trigger the generation of sensory neurons in the dorsal spinal cord from neural crest cells (Sardar, N., et al., 2025). Moreover, by proliferation and differentiation of stem cells, GDF7 also plays a part in the repair and regeneration of ligaments and tendons. Research has shown that it is also involved in supporting the in vitro human adipose-derived stem cell differentiation. Additionally, through regulation of the STING/AMPK pathway, GDF7 makes improvements to sepsis-induced acute lung injury (Alternative Medicine, 2023).

The higher chances of esophageal adenocarcinoma and Barrett's esophagus might be linked to GDF7's genetic alterations. The other diseases associated with GDF7 are Epicondylitis and Shoulder Impingement Syndrome (Raleigh, 2024). Barrett's esophagus known as is a condition in which usual esophageal squamous epithelium gets transformed into metaplastic columnar epithelium, and people with BE have a chance of 2–4% for esophageal adenocarcinoma (EAC), which represents the esophagitis-metaplasia-dysplasia-adenocarcinoma sequence's final stage (Li, Hoefnagel and Krishnadath, 2023). The main trigger for Barrett's esophagus (BE) is chronic gastroesophageal reflux disease (GERD), wherein stomach and bile acids reflux into the esophagus, leading to inflammation and metaplastic changes. Most of the cases are acquired, but rare familial forms suggest a genetic predisposition (Maev *et al.*, 2023).

Typical indications of BE include heartburn, dysphagia, acid regurgitation, and a feeling of a lump within the throat. Its detection in the general population is approximately 1.3 to 1.6%, and it affects 5 to 15% of patients with symptoms of GERD who are undergoing an endoscopy. The occurrence rates vary globally, with higher rates observed in Western countries (Li, Cao and Xu, 2025), where EAC is regarded as a significant disease because of its poor prognosis and is responsible for about 60% of all esophageal cancers(De Melo Viana *et al.*, 2025). Recent studies identified BE interconnection on chromosome 2p24 (at rs3072) and 12q24 (at rs2701108), with GDF7 and TBX5 being implicated as promising risk genes, respectively. Further analysis showed that rs3072 and rs2701108 were also associated with EAC as well, suggesting a continued risk through the BE/EAC sequence (Martinez-Uribe, Becker and Garman, 2026).

The single nucleotide polymorphisms (SNPs) comprise over 90% of variations in sequence in the human genome, and these are prevalent genetic variations that constitute approximately 1% of the entire genome. There are about 3–10 million SNPs, with non-synonymous SNPs (nsSNPs) in coding regions that potentially affect protein function and contribute to various diseases. Computational methods have emerged as efficient and cost-effective tools for screening and predicting the effect of these nsSNPs on structure and function of proteins (Mia *et al.*, 2025a; Saqib et al., 2026). Our study aims to investigate the structural alterations and functional impacts of non-synonymous Single Nucleotide Polymorphisms (nsSNPs) in GDF7 and molecular docking for targeting Barrett's Esophagus.

The current study investigates the association of Barrett's Esophagus **l**inked with the GDF7 gene**,** finding deleterious SNPs and evaluating their impact on the protein structure, function, and interactions through advanced computational analysis. By analyzing large-scale of genetic datasets, this approach efficiently filters and prioritizes potential genetic variants for further experimental validation. The findings offer important insights into the molecular mechanisms underlying Barrett's Esophagus and highlight key genetic contributors, laying a foundation for future wet-lab experiments. The findings offer important insights into the molecular mechanisms underlying Barrett's Esophagus**,** highlighting key genetic factors and providing a strong basis for future wet-lab investigations. These results not only improve the understanding the progression of disease but also support the development of targeted therapeutic interventions for Barrett's Esophagus patients. Moreover, this study supports future population-specific research by helping in the design of wet-lab experiments that consider genetic variability associated with Barrett's Esophagus.

Methodology

Non-Synonymous SNPs' Retrieval

The single-nucleotide polymorphisms (SNPs) of the GDF7 gene were obtained using the database called the NCBI SNP database (dbSNP) (https://www.ncbi.nlm.nih.gov/snp/ and Ensemble database (https://www.ensembl.org/index.html GDF7 gene encodes a protein whose primary sequence (Uniprot accession number: Q7Z4P5) was fetched from the NCBI database (https://www.ncbi.nlm.nih.gov/). From the NCBI SNP database, only non-synonymous SNPs were selected as they can alter the sequence of amino acids and potentially interfere with the form and function of the protein.

Functional Analysis of Non-Synonymous SNPs

Five distinct biological data analysis tools (PolyPhen-2, SIFT, PhD-SNP, SNPs&GO, and Meta-SNP) were utilized to analyze the functional impact of nsSNPs on the protein. In our study, if at least 3 or 4 of these tools indicate that nsSNPs are deleterious, then they are considered critical-risk and undergo further investigation.

At first, the SIFT server (https://sift.bii.a-star.edu.sg/) was utilized to figure out the phenotypic outcomes on the functionality of the protein due to alteration in the amino acid sequence. It categorizes substitutions as tolerated or deleterious at every site in the protein sequence based on the SIFT score (≤ 0.05 were

considered deleterious, and those with a score > 0.05 were classified as tolerated (Jamali, Zargar and Modarressi, 2024). Then PolyPhen-2, a sequence-based mutation analysis tool, was utilized to calculate the potential damaging impacts of a mutation by considering physical and conservative features. An amino acid shift is considered deleterious if the Position-Specific Independent Count (PSIC) score is >0.5 (Chatterjee and Banerjee, 2025). After PolyPhen-2, the SNPs & Go server (https://snps.biofold.org/snps-and-go/snps-and-go.html) was used, which functions to establish the relationship between the disease and the studied SNPs (E. Capriotti *et al.*, 2013). At next, the PhD-SNP server (https://snps.biofold.org/phd-snp/phd-snp) was used, which distinguishes single-point mutations as either neutral polymorphisms of disease-related using the Support Vector Machine (SVM) method (Hossain *et al.*, 2025). Furthermore, the Meta-SNP tool (https://snps.biofold.org/meta-snp/) was used to identify disease-causing variants. It enables the detection of disease-associated nsSNPs (non synonymous single nucleotide variants) for both known and predicted amino acid (Jibon *et al.*, 2025).

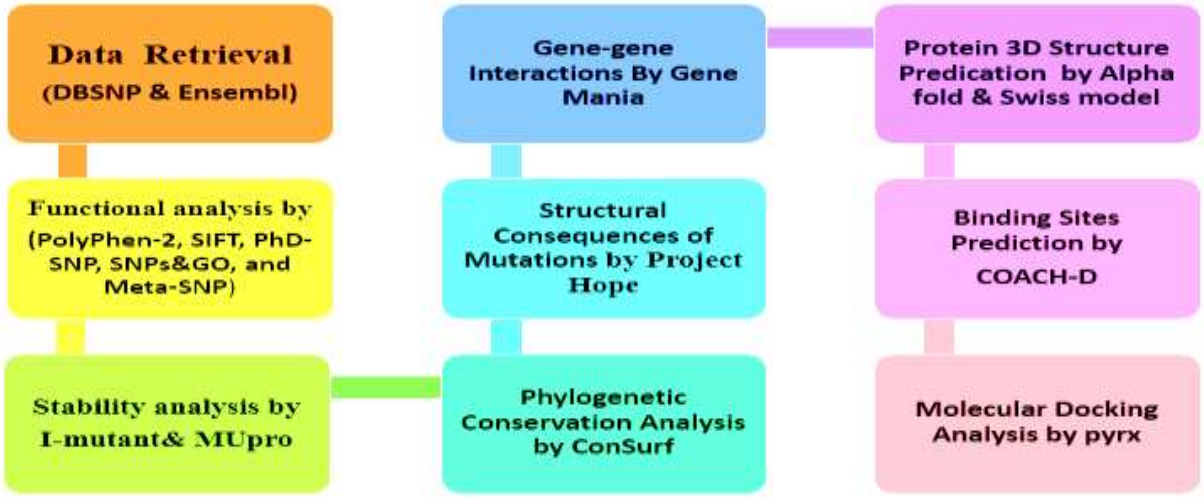


Figure 1: Workflow for screening nsSNPs in the GDF7 gene using computational tools

Protein Stability Analysis

I-Mutant2.0 tool (https://folding.biofold.org/i-mutant/i-mutant2.0.html) was used to automatically figure out the alteration in protein reliability resulting from single-point mutations (E. Capriotti *et al.*, 2008). Then MUpro (https://mupro.proteomics.ics.uci.edu/) was utilized for further forecasting stability of protein alterations resulting from mutations (Akbulut, 2022). It employs support vector machine (SVM) technology to predict these changes, considering structural characteristics or the sequence of the protein based on ΔΔG values (E. Capriotti *et al.*, 2011).

Phylogenetic Conservation Analysis by ConSurf

ConSurf (https://consurf.tau.ac.il/) was used for the identification of functional regions in macromolecules by assessing the evolutionary retention of amino acids or nucleotide substitutions based on evolutionary relationships among sequence homologs. It assigns preservation scores ranging from 1 to 9 in accordance with the evolutionary proportion of sites using empirical Bayesian or maximum likelihood (ML) approaches (Tuya *et al.*, 2024).

Structural Consequences of Mutations

HOPE (http://www.cmbi.ru.nl/hope) was used for analyzing the variation in the structure and function of protein sequences caused by single amino acid substitution. The purpose of this server is to figure out 3D designs of proteins by collecting structural information and sequence tags from the calculations on 3D coordinates of protein and UniProt database, respectively (Shao, Grams and Gao, 2023).

Gene-gene Interactions by Gene MANIA

Gene MANIA (http://www.genemania.org) was used for forecasting gene function and providing knowledge on co-localization, co-expression, shared protein domains, and associated pathways of the gene (Liu, Wang and Li, 2025). It is a user-friendly web tool that helps to generate hypotheses about the function of a gene, analyze gene lists, and prioritize genes for functional assay (Salma *et al.*, 2023).

Protein 3D structure predication

The full length of human GDF7 (Uniprot kb: Q7Z4P5) was obtained through a dual approach. A high quality homology modeling constructs three-dimensional structures from amino acid sequences by employing homologous proteins with known structures as templates (Abul Seoud, El Gali and Abdel Malek, 2024). We utilized the Alpha Fold database to predict the wild-type structure of GDF7 and the Swiss Model to predict mutated models for the selected protein structure corresponding to specific amino acid substitutions.

The Alpha Fold (https://alphafold.ebi.ac.uk/) is a database that provides proposed secondary and tertiary structures for proteins, accessible to the public, facilitating easy access to predicted protein structures even for nonexperts (Varadi *et al.*, 2024). SWISS MODEL (http://swissmodel.expasy.org) is a tool designed for the automated comparative

modeling of three-dimensional (3D) protein structures (Paritala, 2022).

Prediction of Binding Sites for Ligand by COACH-D

The software COACH-D (http://yanglab.nankai.edu.cn/COACH-D/) was used to predict binding sites of the GDF7 protein, and then it docks ligands into these sites to refine complex structures. Benchmark tests demonstrate COACH-D's superiority in accuracy, reducing steric clashes by 85% compared to other methods (Liu *et al.*, 2022).

Molecular Docking Analysis

To evaluate the impact of mutation over the binding affinity GDF7, molecular docking was perform using PyRx (https://pyrx.sourceforge.io/) was is a user-friendly virtual screening software that is used to inspect the binding affinity of multiple ligands against a target protein (Prajapati *et al.*, 2023). The whole protein was encompassed within the grid box during docking. The effectiveness of ligand-target protein interaction is predominantly assessed through the scoring function for that purpose the PDB format of the input ligands were converted into pdqt format (Zulfat *et al.*, 2024). Pymol and Discovery Studios was employed to visualization of the dock complex interactions (Shamsi *et al.*, 2024).

Results

A comprehensive analysis was conducted utilizing various bioinformatics tools and databases, such as Ensembl and dbSNP for the identification of function single nucleotide polymorphisms within the GDF7 gene. This analysis a total 6045 SNPs were retrieve form dbSNP.Further examination of the non-coding region 176 synonymous, and 910 intronic coding region SNPs revealed 325 non-synonymous (missense) in alteration to the encoded amino acid sequence (Figure 2**)** We selected non-synonymous SNPs (nsSNPs) because these possessed the power to affect the protein's structure and function.

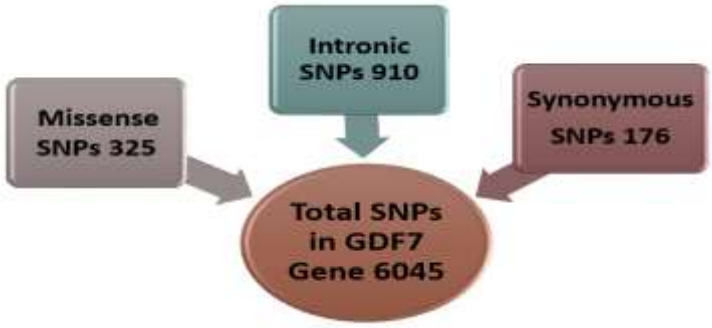


Figure 2: Workflow for SNPs analysis in the GDF7 gene

Prediction of Deleterious Non-Synonymous SNPs

To identify potentially deleterious SNPs that may substantially impact the structure or function of the GDF7 protein, we employed a comprehensive in-silico analysis utilizing six distinct prediction tools: SIFT, PolyPhen-2, Meta-SNP, SNPs&GO and polyphen-2. Our analysis revealed that four non-synonymous SNPs (nsSNPs) consistently exhibited deleterious predictions across all five computational methods, suggesting their potential to induce harmful effects on protein function. At first, among 325 nsSNPs, 9 were found to be deleterious by SIFT, having a score ≤ 0.05. The reliability of prediction was confirmed when these 9 nsSNPs were also found to be deleterious by PolyPhen2, having a score > 0.90. Then the other bioinformatics tools, i.e., SNPs&GO, PhD-SNP, and Meta SNP, were used for further confirmation (Table 1), and seven deleterious nsSNPs were selected for further analysis.

Table 1: **Deleterious non-synonymous SNPs in GDF7 gene**

rs ID	Amino Acid Change	SIFT	Poly-phen 2	SNPs & GO	phD-SNPs	Meta-SNP
rs142632106	D429N	Deleterious	Probably Damaging	Disease	Disease	Disease
rs146317060	D364H	Deleterious	Probably Damaging	Disease	Neutral	Disease
rs146567848	R419H	Deleterious	Probably Damaging	Neutral	Disease	Disease
rs184953707	L234M	Deleterious	Probably Damaging	Neutral	Neutral	Disease
rs367794645	K352T	Deleterious	Probably Damaging	Neutral	Neutral	Disease
rs369544272	V434A	Deleterious	Probably Damaging	Disease	Neutral	Disease

rs377143734	I423M	Deleterious	Probably Damaging	Disease	Neutral	Disease

Stability Analysis of Non-Synonymous SNPs

To evaluate the potential of the prioritized nsSNPs harmful mutation, we conducted a computational analysis of their impact on the GDF7 protein stability and function. Five nsSNPs, rs146317060, rs146567848, rs367794645, rs369544272, rs377143734, were found to decrease the stability of the protein by the I-Mutant 2.0 server. However, MUpro prediction was different. A negative score from the MUpro server signifies a decrease in protein structural stability following a substitution, while a positive score indicates an increase. By comparing the results of both servers, a decrease was found in the stability of these five nsSNPs (Table 2). Based on the comparison, these 5 deleterious nsSNPs were selected to proceed towards gene-gene interaction, conservation analysis, and structural consequences prediction by using Gene Mania, ConSurf, and Project Hope, respectively.

Table 2: Protein stability analysis of nsSNPs in GDF7

rs ID	Amino Acid Change	I-Mutant Prediction	MUpro Prediction
rs146317060	D364H	Decrease P/H	Decrease Stability
rs146567848	R419H	Decrease D/H	Decrease Stability
rs367794645	K352T	Decrease E/T	Decrease Stability
rs369544272	V434A	Decrease G/A	Decrease Stability
rs377143734	I423M	Decrease A/M	Decrease Stability

Consurf's Conservational Analysis

The amino acid residues evolutionary conservation in the native GDF7 protein was analyzed using the ConSurf web server. ConSurf analysis assessed the significance of 5 high-risk nsSNPs structures and functions while based on evolutionary conservation and solvent accessibility. The results indicated that residues D364, K352, V434, I423, and R419, functionally important and exposed, are structurally significant and buried. Notably, all 5 of these residues exhibit high conservation. Consurf predicted that our selected nsSNPs were located in conserved regions, while D364H, V434A, and I423M were highly conserved (Figure 3).

Figure 3: Evolutionary Conservation Mapping of the GDF7 Protein. The conservation pattern of amino acid residues is represented using nine-grade color Scale highlighting highly conserved and variable region

Gene-to-Gene Interaction Predicted by Gene Mania

A protein interaction network was constructed for GDF7, revealing its functional associations with multiple gene or protein. These gene were grouped into distinct clusters, with GDF7 and BMP7 acting as central hubs. The network includes key signaling molecules such as ACVRL1, ACVR1, ACVR1B, ACVR1C, ACVR2A, ACVR2B, AMHR2, CHRD, NOG, BMP10, BMP7, BMPR1A, BMPR1B, BMPR2, GDF6, TRH, PAMR1, BAMBI, SORBS1, and FKBP1A. Any structural or functional alterations in GDF7 could potentially disrupt interactions within this network, affecting the stability and regulatory roles of the associated proteins. Gene MANIA generates a network of functional interconnection among genes, and the network of interaction of the GDF7 gene is presented in Figure 4.

Comparative Structural Analysis of Mutant and Wild-Type GDF7 Models

We utilized the Swiss model for homology modeling tool and AlphaFold to generate the 3D structures of both the native GDF7 protein and its mutant variants. The wild-type GDF7 structure is predicted by alpha Fold, and the mutant GDF7 structure is predicted by SWISS-MODEL. Each of the identified deleterious nsSNPs was individually introduced into the mutant sequence of the respective templates, and the 3D structural models of all mutants were subsequently predicted. Five mutations (D364H, R419H, K352T, V434A, and I423M) were incorporated in the modeling of the mutant structure, which were confirmed by multiple software tools as deleterious. SWISS-MODEL resulted in 6 templates, of which one template was Q9BDW8.1.A Growth/differentiation factor 7 was found to be most suitable, with 94.63% identity with the amino acid sequence provided (Figure 5).

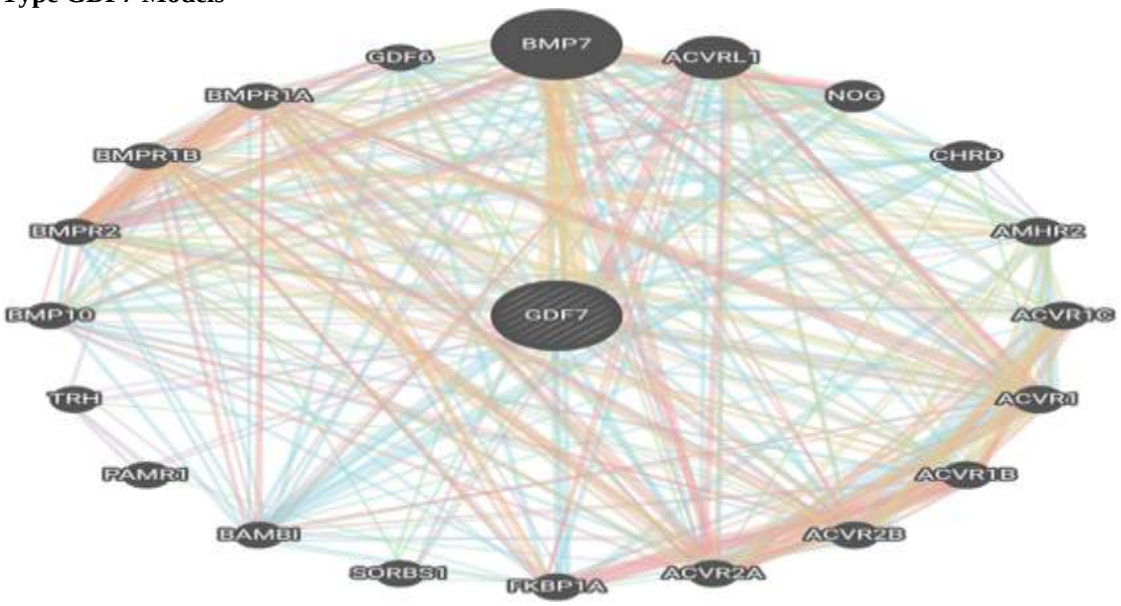


Figure 4: Gene-gene interactions obtained by Gene Mania

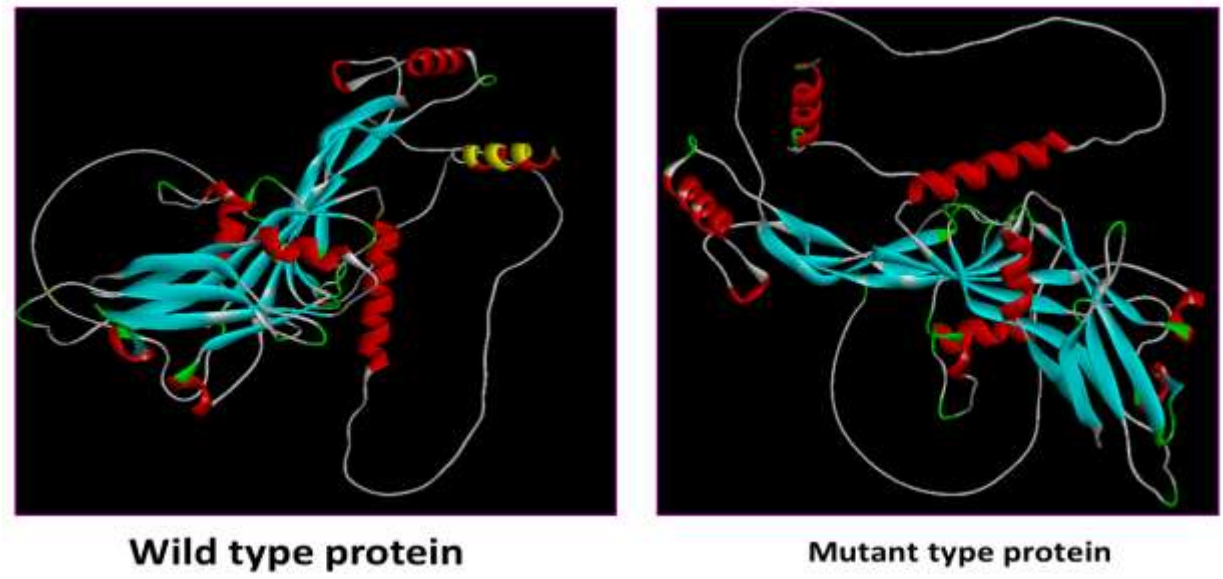
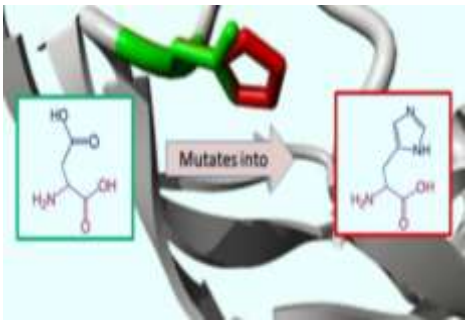
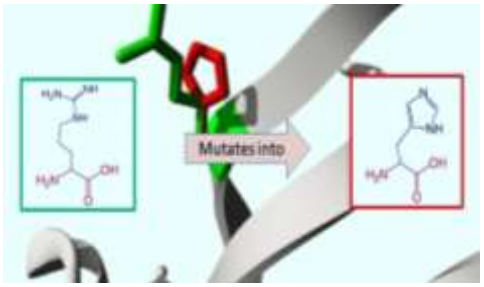
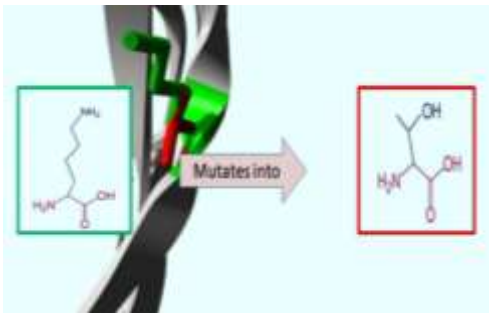
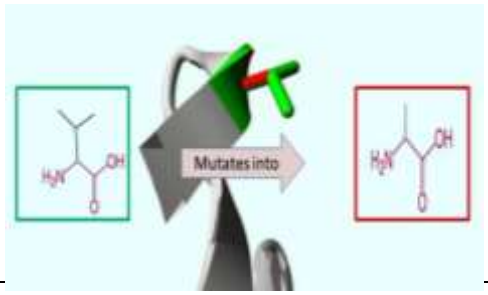


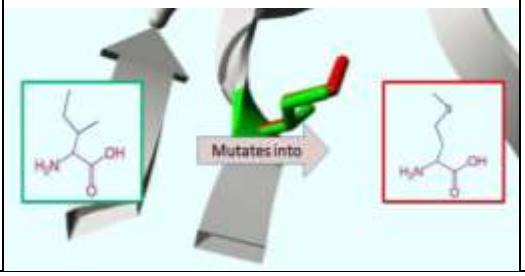
Figure 5: Wild-type and mutant peptide of GDF7 protein

Structural Impact of Point Mutations on GDF7 Protein

The HOPE tool used to analyze the effects of variation on the LYZ, 3D structure and function through comparing the physicochemical properties between wild-type and variant amino acids. Effect of these variations has been shown in Table 3. The Project HOPE server analysis revealed that the mutant residues are larger and more hydrophobic than the wild-type residues. This change in hydrophobicity and size may disrupt the hydrogen bond interactions with adjacent molecules.

Table 3 Hope predictions of structural consequences of mutations

rsID	Amino Acid Change	Effect of Substitution
rs146317060	**D364H** 	Mutant residue has been found bigger as compared with wild-type residue. Negative charge of occurred in wild-type residue was lost due to this mutation, which became the cause for the loss of interactions for other molecules.
rs146567848	**R419H** 	The mutant residue has been found smaller as compared with wild-type residue. Positive charge of wild-type residue is lost due to this mutation. It may become the cause for possible loss for external interactions.
rs367794645	**K352T** 	The mutant residue has been found smaller and more hydrophobic as compared with the wild-type residue. Positive charge of wild-type residue lost due to this mutation, which may result in a loss for the interactions with other molecules.
rs369544272	**V434A** 	The mutant residue is smaller as compared with the wild-type residue, which results in a possible loss for the external interactions.

rs377143734	**I423M** 	The wild-type residue was found to be buried in the core of the protein. The mutant residue is bigger as compared with the wild-type residue which probably will not fit.

Binding Sites' Prediction

Binding site prediction is a crucial computational approach for identifying potential ligand-binding regions within a protein structure. The COACH server analysis revealed that both wild-type and mutant structures exhibit 122 to 450 amino acid residues involved in binding site formation. The comparison between wild-type and mutant proteins showed alterations in key binding residues. The Coach server predicted binding sites of the GDF7 protein, and then it docked ligands into these sites to refine complex structures, as shown in Figure 6.

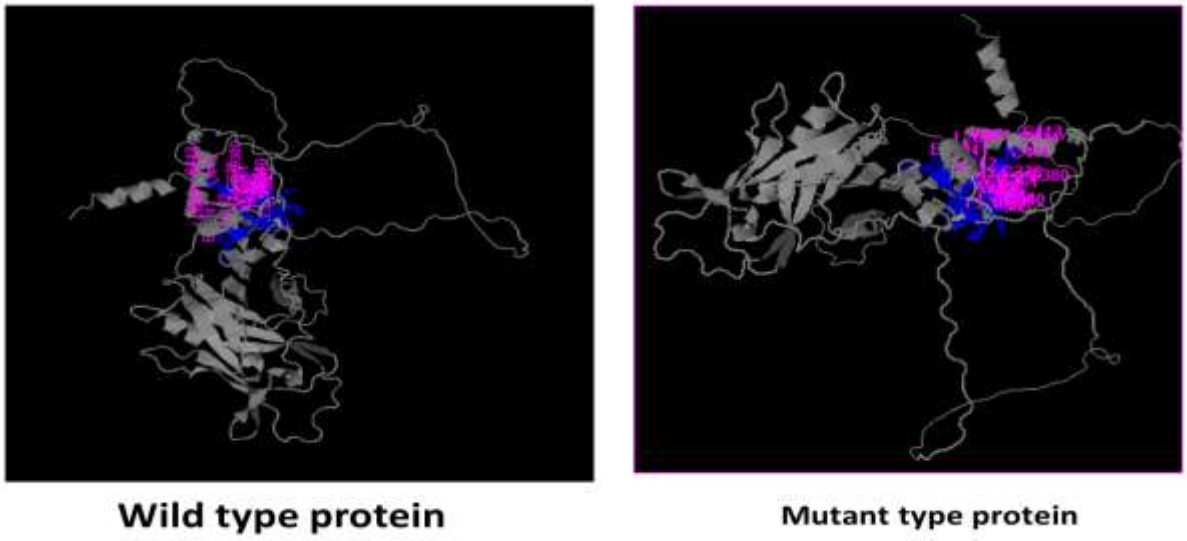


Figure 6: Coach-D results of binding regions determined at different amino acid positions

Docking Analysis

The molecular docking findings evaluated the binding affinity for the mutated GDF7 protein and wild-type GDF7 protein, as shown in Figure 7. Among the mutant proteins analyzed, showed a significant reduction in binding affinity. The interaction patterns and binding affinity (kcal/mol) of these docked complexes were analyzed through Discovery Studios and Pymol software. The wild-type peptide sequence formed an affinity of -7.1 kcal/mol, whereas the mutant formed only, reducing binding affinity to -6.6 kcal/mol.

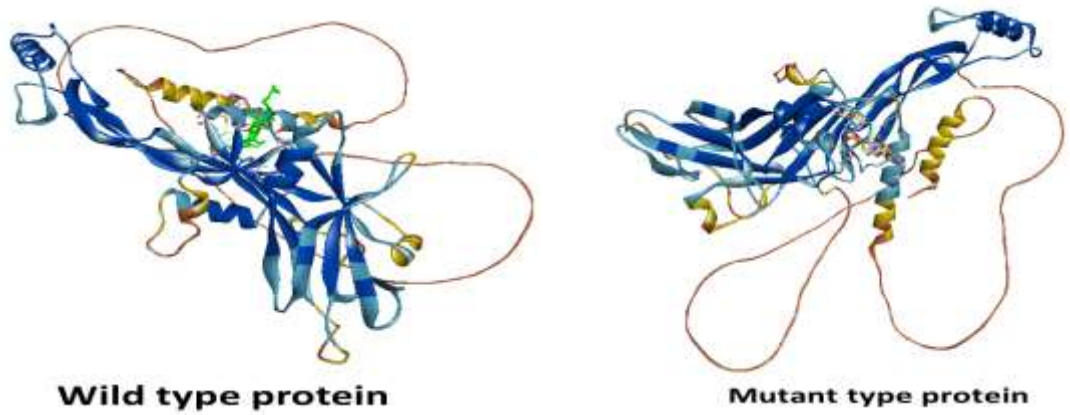


Figure 7: Molecular docking analysis of the wild-type and mutant peptide of GDF7 protein

The binding interactions are further illustrated by a 2D interaction diagram as shown in Figure 8 and Table 4. For the wild-type GDF7, the significant interactions are van der Waals forces, hydrogen bonds, Pi-Alkyl, and Alkyl interactions, and eight key residues are involved, i.e., ARG221, SER272, ASP150, LEU148, ASN149, SER147, VAL270, PHE124, and VAL145. While in the mutant GDF7, different interactions are observed i.e. Pi-Anion and Pi-Donor Hydrogen Bonds and four key residues are involved i.e. GLU283, ASP429, TYR427, and VAL435., SER 147, ALA 290. In the mutant structure, Pi-Anion and Pi-Donor interactions are unique, and these present altered electronic as well as stable interactions.

.

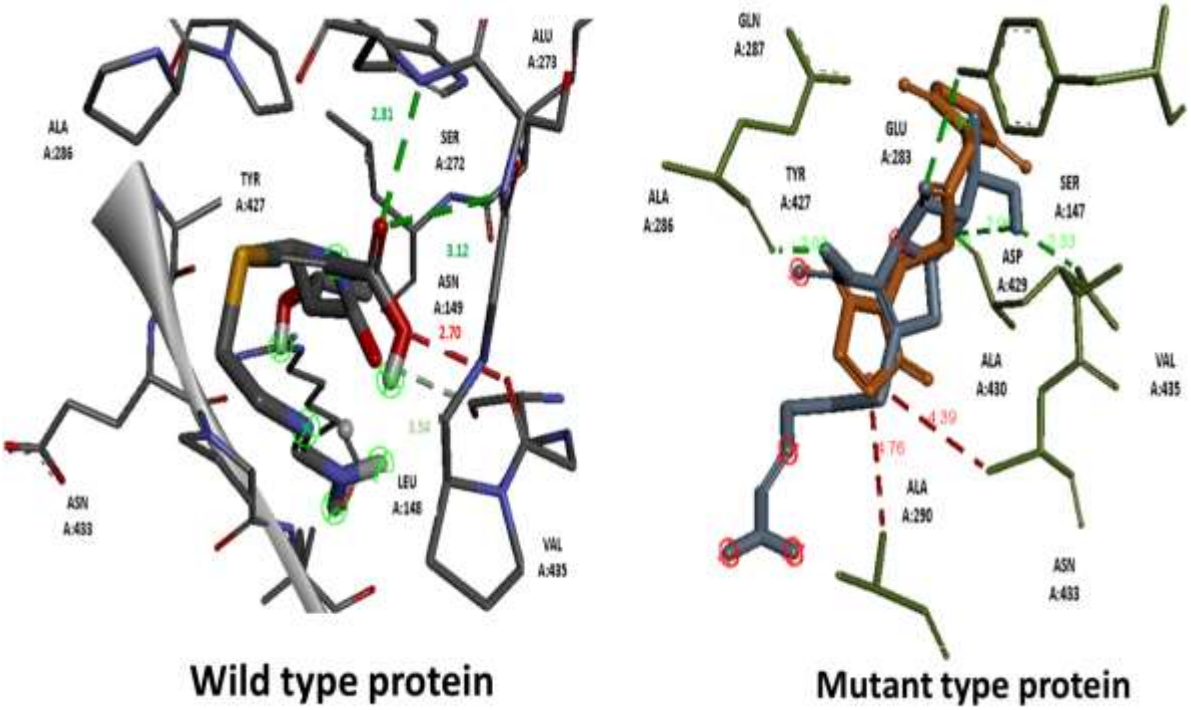


Figure 8: Structural impact of the GDF7 mutation. (a) Wild-type protein hydrogen bond analysis (b) Mutant-type protein hydrogen bond analysis

Table 4: The molecular docking results show the binding affinities and interaction scores of the compounds with the target protein

Protein Type	Binding Affinity	Chimera Analysis of Interactions
Wild-type protein	-7.1 kcal/mol	Van der Waals interactions with ARG221, , ASP150, LEU 148, ASN149, SER147. Pi-Alkyl interactions with SER272, PHE124, VAL145. Conventional Hydrogen Bond with ASN149. Alkyl interactions with SER272.

Mutant-type protein	-6.6 kcal/mol	Conventional Hydrogen Bond with GLU283 and ASP429. Pi-Anion interactions with GLU283. Pi-Donor Hydrogen Bond with TYR427. Pi-Alkyl interactions with VAL435.

Discussion

GDF7 plays a critical function in cell survival, differentiation, and development, particularly in the case of neural and tendon tissues, emphasizing its significance in tissue repair and embryonic development through the activation of SMAD transcription factors (Lee *et al.*, 2024). Previous studies have established a connection between GDF7 and Barrett's esophagus (BE), suggesting its role in case of gastrointestinal disorders (Yang et al., 2026).). The 325 nsSNPs have been reported in the GDF7 protein, but there is a research gap in understanding their implications, disease connections, phylogenetic conservation, or effects on protein stability and functionality. This study was proposed to address this gap by employing various *In-silico* tools for analysis of the effects and associations of nsSNPs on functionality and stability of GDF7. To predict harmful nsSNPs of GDF7, web-based technologies such as SIFT, PolyPhen 2, SNPs&GO, PhD-SNP, Meta SNP, I-mutant, MUpro, Gene MANIA, Consurf, and Project Hope were used. Initially, SIFT and PolyPhen were utilized to assess deleterious nsSNPs of the GDF7, which were then further inspected using PhD-SNP, SNPs&GO, and Meta SNP. Seven nsSNPs were identified as deleterious by merging the results which are rs142632106 (D429N), rs146317060 (D364H), rs146567848 (R419H), rs184953707 (L234M), rs367794645 (K352T), rs369544272 (V434A), rs377143734 (I423M). Changes in protein stability serve as a recognized mechanism through which amino acid substitutions contribute to human diseases, and the majority (83%) of nsSNPs that are disease-causing have the potential to influence the stability of the protein (Shinwari *et al.*, 2023). A protein's structural stability is important for adequate activity level, regulation, and function. Functional impact may arise from protein misfolding, degradation, and accumulation caused by a decline in stability. A protein's stability is reflected in its folding free energy (ΔG), which evaluates the energy difference between its folded and unfolded forms. The difference between the ΔG values of wild-type and mutant-type proteins is known as the folding free energy change (ΔΔG) (Ajmal, 2023). Protein stability is increased by values greater than 0 and decreased by values less than 0. I-Mutant and MUpro were used to assess the effect of detrimental nsSNPs on the stability of proteins. Subsequently, these 7 nsSNPs underwent further analysis by I-Mutant and MUpro servers, which determined that nsSNPs—rs146317060, rs146567848, rs367794645, rs369544272, rs377143734—decreased protein stability. Each predicted nsSNPs decreased the stability of the protein, which may affect its functionality. These results provide a potential mechanistic explanation for previously discovered associations between these gene variations and disease characteristics.

The gene interaction network of the gdf7 gene is linked to multiple additional genes involved in various kinds of disorders, as predicted by GENEMANIA.GeneMANIA shows interactions between GDF7 and related genes through physical interaction approximately 77%, co-expression approximately 8%, predicted approximately 5%, co-localization approximately 3.6%, genetic interactions approximately 2.8%, pathways approximately 1.8%, and shared protein domains approximately 0.6%. Analyzing evolutionary conservation for a protein residue helps identify potential harmful effects of mutations on the host. Genetic variations among interacting genes can affect GDF7's biological activity and functional regulation, which may increase the development of several types of diseases. As a consequence, analyzing these networks offers important new perspectives on the molecular processes and functional connections that characterize complicated biological systems(Velmurugan *et al.*, 2026). The conserved region of amino acid residues has a significant impact on the structural and functional characteristics of proteins during evolution. Consequently, nsSNPs occurring at these important conserved positions are considered more likely to exert harmful effects than variants present in variable or less conserved regions.ConSurf's result shows that our selected 5 mutations were found in conserved locations, and D364H, V434A, and I423M were highly conserved. These 5 deleterious nsSNPs, i.e., D364H, R419H, K352T, V434A, and I423M, went through additional *In-silico* assessment.

To determine the effect of substitution of amino acids at the molecular level, the Project Hope was used, which shows changes in hydrophobicity, size, and charge. Moreover, the pathogenicity of our SNPs was confirmed because all the SNPs were positioned in the Transforming Growth Factor-Beta, C-Terminal IPR001839 domain. D364H and I423M were predicted to be larger, and R419H, K352T, and V434A were predicted to be shorter, which may cause a decline in interactions with other molecules. The interaction may also be lost because of getting a neutral charge in case of mutations, i.e., D364H, R419H, K352T. The mutation, i.e., K352T, may cause a decrease in hydrogen bonds, as its hydrophobicity was found to be different (Aleshina, Zavileyskiy and Aleshin, 2025).

The investigation was further extended to analyze the structural consequences of these deleterious nsSNPs using homology modeling tools. Based on the selected templates, wild-type and mutant protein models were

generated for the GDF7 protein(Bahereldeen *et al.*, 2020). The structural prediction was carried out using homology modeling and AlphaFold for the wild-type structure (Figure 5), while SWISS-MODEL was used to model the mutant structure (Figure 6), incorporating five deleterious mutations (D364H, R419H, K352T, V434A, and I423M). A total of six templates were identified by SWISS-MODEL. Among these, template Q9BDW8.1.A, corresponding to Growth Differentiation Factor 7, was selected as the most suitable due to its high sequence identity of 94.63% with the query amino acid sequence. This high sequence identity ensures that the modeled mutant structure closely resembles the native protein, thereby providing a reliable basis for subsequent docking and interaction analyses. Furthermore, the protein binding regions were predicted using COACH-D (Table 4). The binding sites can facilitate the design of small molecules or drugs targeting the GDF7 protein for therapeutic intervention in disease. Similarly, these results align with previous studies reporting that deleterious nsSNPs can significantly affect protein structure and function, particularly in growth-related proteins. Collectively, previous research supports that mutations could lead to changes in protein function, structure, and stability that contribute to disease-related pathways (Vallejos-Vidal *et al.*, 2020).

Computational methods enable the initial evaluation of compounds' binding potential for therapy, with molecular docking widely used in drug designing and compound screening to assess their affinity for target proteins (Naithani and Guleria, 2024). In this study, Proflavine was selected as the compound for docking with GDF7 using PyRx. A score is assigned by the scoring function on the best docked ligand complex (Zhou, Li and Chen, 2025). In case of molecular docking, binding affinities that are less than -6 kcal/mol are typically considered to represent strong binding, whilst those that are roughly -5 kcal/mol are considered to represent moderate binding (Xu *et al.*, 2020). The binding affinity values for normal GDF7, i.e., -7.1 kcal/mol, and for mutant GDF7, i.e., -6.6 kcal/mol, suggest strong binding interactions. This is consistent with previous studies that have shown that nsSNPs can affect protein function by altering binding affinities .However, it is important to note that experimental validation is necessary to confirm the results of the docking as the mutations may have different effects on the binding of other ligands

The Docking analysis with Chimera revealed significant interactions between Proflavine and both wild-type and mutant GDF7 proteins. In the wild-type GDF7 protein, hydrogen bonds with ASN149 and van der Waals forces were found along with stabilizing Pi-Alkyl and Alkyl interactions, and this suggests strong and stable binding. In the mutant GDF7 protein, in addition to conventional hydrogen bonds, different types of interactions like Pi-Anion and Pi-Donor Hydrogen Bonds are observed, and 4 residues are involved, which depicts an altered binding site environment. According to the docking result, as Proflavine has strong binding affinity and a variety of interactions towards mutated GDF7, it could be a promising candidate for targeting the mutated pathway of the gene. Overall, these findings demonstrate that docking-based approaches are effective in elucidating mutation-induced changes in protein structure and stability (Hossain, Roy and Islam, 2020). Due to non-synonymous mutations in GDF7, the normal signaling pathway may be disrupted. Such alterations may consequently contribute to abnormal signaling behavior and disease-associated molecular changes. So, this study will provide a substantial framework for the identification of functional SNPs.

Conclusion

In conclusion, our extensive *in-silico* analysis elucidates the impact of that nsSNPs GDF7 gene significantly impact its structure and function, potentially contributing to Barrett's Esophagus. Our findings suggest that alterations in the GDF7 gene may influence patient prognosis, highlighting the need for further experimental validation and functional studies. The identification of these nsSNPs presents potential targets within the GDF7 pathway for therapeutic interventions, paving the way for novel treatment strategies for this progressive disease. This study contributes significantly to our understanding of Barrett's Esophagus genetic basis and underscores the importance of considering nsSNPs functional effects in GDF7, ultimately informing diagnosis, prognosis, and treatment approaches.

References

Abul Seoud, R., El Gali, N., & Abdel Malek, M. (2024). Protein 3D structure prediction using homology method. *Fayoum University Journal of Engineering*, ***7***(2), 338–346. https://doi.org/10.21608/fuje.2024.345050

Ajmal, M. R. (2023). Protein misfolding and aggregation in proteinopathies: Causes, mechanism and cellular response. *Diseases*, ***11***(1), Article 30. https://doi.org/10.3390/diseases11010030

Akbulut, E. (2022). Investigation of changes in protein stability and substrate affinity of 3CL-protease of SARS-CoV-2 caused by mutations. *Genetics and Molecular Biology*, ***45***(3), Article e20210404. https://doi.org/10.1590/1678-4685-gmb-2021-0404

Aleshina, Y. A., Zavileyskiy, L. G., & Aleshin, V. A. (2025). Neanderthal and Denisovan glutamate dehydrogenase 2 evolution and clinical significance. *International Journal of Molecular Sciences*, ***26***(9), Article 4322. https://doi.org/10.3390/ijms26094322

Evidence-Based Complementary and Alternative Medicine. (2023). Retracted: Growth differentiation factor 7 prevents sepsis-induced acute lung injury in mice. *Evidence-Based*

Complementary and Alternative Medicine, ***2023***(1), Article 9795462. https://doi.org/10.1155/2023/9795462

Bahereldeen, A. I., et al. (2020). *In silico analysis of single nucleotide polymorphisms (SNPs) in human HFE gene coding region*. [Unpublished manuscript / Preprint].

Chatterjee, P., & Banerjee, S. (2025). A computational and structural approach to identify malignant non-synonymous FOXM1 single nucleotide polymorphisms in triple-negative breast cancer. *Scientific Reports*, ***15***(1), Article 964. https://doi.org/10.1038/s41598-024-85100-w

E. Capriotti, P. Fariselli, I. Rossi, and R. Casadio, 'A three-state prediction of single point mutations on protein stability changes', BMC Bioinformatics, vol. 9, no. SUPPL. 2, pp. 1–9, 2008, doi: 10.1186/1471-2105-9-S2-S6.

E. Capriotti and R. B. Altman, 'Improving the prediction of disease-related variants using protein three-dimensional structure', BMC Bioinformatics, vol. 12, no. SUPPL. 4, 2011, doi: 10.1186/1471-2105-12-S4-S3.

De Melo Viana, T. C., et al. (2025). Molecular abnormalities and carcinogenesis in Barrett's esophagus: Implications for cancer treatment and prevention. *Genes*, ***16***(3), Article 270. https://doi.org/10.3390/genes16030270

Hossain, Md. M., et al. (2025). Computational prediction of high-risk non-synonymous SNPs in human ApoE and their structural impact on amyloid-β interaction in Alzheimer's disease pathogenesis. *PLOS ONE*, ***20***(9), Article e0331339. https://doi.org/10.1371/journal.pone.0331339

Hossain, Md. S., Roy, A. S., & Islam, Md. S. (2020). In silico analysis predicting effects of deleterious SNPs of human RASSF5 gene on its structure and functions. *Scientific Reports*, ***10***(1), Article 14542. https://doi.org/10.1038/s41598-020-71457-1

Ishaq, S., Ali, A., Habib, O., Usmani, S. A., Rehman, I. U., Aziz, A., & Ullah, A. (2026). Pathogenic KRAS variants disrupt structure and dynamics: Insights from integrated computational analyses. *PLOS ONE*, ***21***(2), Article e0341219. https://doi.org/10.1371/journal.pone.0341219

Jamali, Z., Zargar, M., & Modarressi, M. H. (2024). Prediction the functional impacts of highly deleterious non-synonymous variants of TSGA10 gene. *Molecular Biology Research Communications*, Online First. https://doi.org/10.22099/mbrc.2024.49991.1977

Jibon, Md. D. K., et al. (2025). In-silico analysis of deleterious non-synonymous SNPs in the human AVPR1a gene linked to autism. *BMC Genomics*, ***26***(1), Article 492. https://doi.org/10.1186/s12864-025-11655-1

Lee, J., et al. (2024). Roles of SMAD and SMAD-associated signaling pathways in nerve regeneration following peripheral nerve injury: A narrative literature review. *Current Issues in Molecular Biology*, ***46***(7), 7769–7781. https://doi.org/10.3390/cimb46070460

Li, M., Cao, S., & Xu, R. (2025). Global trends and epidemiological shifts in gastrointestinal cancers: Insights from the past four decades. *Cancer Communications*, ***45***(7), 774–788. https://doi.org/10.1002/cac2.70017

Li, S., Hoefnagel, S. J. M., & Krishnadath, K. K. (2023). Molecular biology and clinical management of esophageal adenocarcinoma. *Cancers*, ***15***(22), Article 5410. https://doi.org/10.3390/cancers15225410

Liu, M., Wang, T., & Li, M. (2025). Identification of immune infiltration-associated CC motif chemokine ligands as biomarkers and targets for colorectal cancer prevention and immunotherapy. *International Journal of Molecular Sciences*, ***26***(2), Article 625. https://doi.org/10.3390/ijms26020625

Liu, Y., et al. (2022). CB-Dock2: Improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Research*, ***50***(W1), W159–W164. https://doi.org/10.1093/nar/gkac394

Maev, I. V., et al. (2023). Esophageal mucosal resistance in reflux esophagitis: What we have learned so far and what remains to be learned. *Diagnostics*, ***13***(16), Article 2664. https://doi.org/10.3390/diagnostics13162664

Martinez-Uribe, O., Becker, T. C., & Garman, K. S. (2026). Pathogenesis of Barrett's esophagus: Evolving and emerging mechanisms. *Advances in Barrett's Esophagus*, ***36***(1), 1–19. https://doi.org/10.1016/j.giec.2025.09.002

Naithani, U., & Guleria, V. (2024). Integrative computational approaches for discovery and evaluation of lead compound for drug design. *Frontiers in Drug Discovery*, ***4***, Article 1362456. https://doi.org/10.3389/fddsv.2024.1362456

Paritala, V. (2022). Comparative modeling for *Saccharomyces cerevisiae* RAD56 using SWISS-MODEL, I-TASSER, and PHYRE2. *International Journal of Pharmacy and Pharmaceutical Sciences*, ***14***(4), 28–33. https://doi.org/10.22159/ijpps.2022v14i4.43972

Prajapati, P. M., et al. (2023). Molecular docking: An advance bioinformatics strategy for structure-based drug designing. In S. Bano et al. (Eds.), *IIP Series* (pp. 184–196). Iterative International Publishers. https://doi.org/10.58532/V3BJBT4P2CH2

Raleigh, S. M. (2024). Genetic and epigenetic factors that predispose to musculoskeletal disorders. *Genes*, ***15***(9), Article 1194. https://doi.org/10.3390/genes15091194

Salma, M., et al. (2023). Var|Decrypt: A novel and user-friendly tool to explore and prioritize variants in whole-exome sequencing data. *Epigenetics &*

Chromatin, ***16***(1), Article 23. https://doi.org/10.1186/s13072-023-00497-4

Sardar, N., et al. (n.d.). *Exploring the impact of STAT4 non-synonymous mutations on Hepatitis B virus susceptibility: A bioinformatics approach.Int.j.mol.microbio.,8(1):79-102* .

Shamsi, A., et al. (2024). Structure-based drug-development study against fibroblast growth factor receptor 2: Molecular docking and molecular dynamics simulation approaches. *Scientific Reports*, ***14***(1), Article 19439. https://doi.org/10.1038/s41598-024-69850-1

Shao, X., Grams, C., & Gao, Y. (2023). Sequence Coverage Visualizer: A web application for protein sequence coverage 3D visualization. *Journal of Proteome Research*, ***22***(2), 343–349. https://doi.org/10.1021/acs.jproteome.2c00358

Shinwari, K., et al. (2023). In-silico assessment of high-risk non-synonymous SNPs in ADAMTS3 gene associated with Hennekam syndrome and their impact on protein stability and function. *BMC Bioinformatics*, ***24***(1), Article 251. https://doi.org/10.1186/s12859-023-05361-6

Tao, H., et al. (2023). Growth differentiation factor 7 pretreatment enhances the therapeutic capacity of bone marrow-derived mesenchymal stromal cells against cerebral ischemia-reperfusion injury. *Chemico-Biological Interactions*, ***386***, Article 110779. https://doi.org/10.1016/j.cbi.2023.110779

Tuya, F., et al. (2024). Strong phylogenetic signal and models of trait evolution evidence phylogenetic niche conservatism for seagrasses. *Journal of Ecology*, ***112***(3), 446–460. https://doi.org/10.1111/1365-2745.14232

Vallejos-Vidal, E., et al. (2020). Single-nucleotide polymorphisms (SNP) mining and their effect on the tridimensional protein structure prediction in a set of immunity-related expressed sequence tags (EST) in Atlantic Salmon (*Salmo salar*). *Frontiers in Genetics*, ***10***, Article 1406. https://doi.org/10.3389/fgene.2019.01406

. https://doi.org/10.1093/nar/gkad1011

Velmurugan, S., et al. (2026). Deleterious nonsynonymous SNPs of human VDR and BMP2 genes and their impact on ankylosing spondylitis: A comprehensive in silico analysis. *Egyptian Journal of Medical Human Genetics*, ***27***(1), Article 5. https://doi.org/10.1186/s43042-025-00834-0

Wiegering, A., et al. (2026). The Joubert gene TMEM67 is required for the correct establishment of spinal dorsal identities in human organoids. *bioRxiv*. https://doi.org/10.1101/2026.01.20.700675

Xu, X., et al. (2020). Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China Life Sciences*, ***63***(3), 457–460. https://doi.org/10.1007/s11427-020-1637-5

Yang, J., et al. (2026). Signaling pathways and targeted interventions for precancers. *Signal Transduction and Targeted Therapy*, ***11***(1), Article 9. https://doi.org/10.1038/s41392-025-02342-4

Zhou, Y., Li, W., & Chen, S.-J. (2025). RLDOCKScore: A scoring function for RNA-ligand docking and small molecule virtual screening. *Journal of Chemical Theory and Computation*, ***21***(19), 9955–9966. https://doi.org/10.1021/acs.jctc.5c00981

Zulfat, M., et al. (2024). Identification of novel NLRP3 inhibitors as therapeutic options for epilepsy by machine learning-based virtual screening, molecular docking and biomolecular simulation studies. *Heliyon*, ***10***(15), Article e34410. https://doi.org/10.1016/j.heliyon.2024.e34410

Declaration

Acknowledgments

We thank the International Islamic University for providing computer lab facilities. The authors are thankful to all the authors who helped and supported in paper proofreading.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

RB: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing**. SR:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Funding acquisition**. SK:** Investigation, Data curation, Formal analysis, Visualization. **NB:** Software, Formal analysis, Data curation. **RS & SM:** Investigation, Resources, Validation. **SK:** Funding acquisition, Writing – review & editing. **AU, QA:** Resources, Project administration, Supervision, Writing – review & editing. **AU:** Conceptualization, Supervision, Writing – review & editing.

All authors read and approved the final manuscript.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Consent to participate

Not Applicable.

Consent to publish

Not Applicable

Ethical Approval and Consent to Participate

Ethical approval is not needed for this review article.

Funding statement

No specific funding or grants were utilized for the accomplishment of this review work.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work, the author(s) have not used any ChatGPT (OpenAI) tool.