

RESEARCH

Open Access

In silico analysis of non-synonymous single nucleotide polymorphisms of human *DEFB1* gene



Harini Venkata Subbiah¹, Polani Ramesh Babu² and Usha Subbiah^{1*} 

Abstract

Background: Single nucleotide polymorphisms (SNPs) play a significant role in differences in individual's susceptibility to diseases, and it is imperative to differentiate potentially harmful SNPs from neutral ones. Defensins are small cationic antimicrobial peptides that serve as antimicrobial and immunomodulatory molecules, and SNPs in β -defensin 1 (*DEFB1* gene) have been associated with several diseases. In this study, we have determined deleterious SNPs of the *DEFB1* gene that can affect the susceptibility to diseases by using different computational methods. Non-synonymous SNPs (nsSNPs) of the *DEFB1* gene that have the ability to affect protein structure and functions were determined by several in silico tools—SIFT, PolyPhen v2, PROVEAN, SNAP, PhD-SNP, and SNPs&GO. Then, nsSNPs identified to be potentially deleterious were further analyzed by I-Mutant and ConSurf. Post-translational modifications mediated by nsSNPs were predicted by ModPred, and gene-gene interaction was studied by GeneMANIA. Finally, nsSNPs were submitted to Project HOPE analysis.

Results: Ten nsSNPs of the *DEFB1* gene were found to be potentially deleterious: rs1800968, rs55874920, rs56270143, rs140503947, rs145468425, rs146603349, rs199581284, rs201260899, rs371897938, rs376876621. I-Mutant server showed that nsSNPs rs140503947 and rs146603349 decreased stability of the protein, and ConSurf analysis revealed that SNPs were located in conserved regions. The physiochemical properties of the polymorphic amino acid residues and their effect on structure were determined by Project HOPE.

Conclusion: This study has determined high-risk deleterious nsSNPs of β -defensin 1 and could increase the knowledge of nsSNPs towards the impact of mutations on structure and functions mediated by β -defensin 1 protein.

Keywords: Defensin, Missense, Polymorphism, In silico analysis

Background

Genetic factors like single nucleotide polymorphisms (SNPs) are one of the most common types of genomic sequence variations and potentially alter disease outcomes. Over a million SNPs are known and lie in the coding regions or within introns and intergenic sequences of DNA that are not coded/translated into

amino acids directly. Missense non-synonymous SNPs (nsSNPs) are of particular interest as a single nucleotide change codes for different amino acids and may affect functions of the encoded protein and disease outcome. The increasing interest in unraveling the relationship between variation in genetic sequences and its effect on a protein's structure and function has led to several *in silico* methods. Experimental techniques to identify the effect of multiple nsSNPs are high-cost, laborious, and involve lots of time, and therefore, in silico approaches can be used as preliminary tools to analyze the effect of nsSNPs.

* Correspondence: ushat75@yahoo.com

¹Human Genetics Research Centre, Sree Balaji Dental College and Hospital, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India

Full list of author information is available at the end of the article

Antimicrobial peptides (AMPs) are important defense peptides of the host and act at the frontline of innate responses displaying a broad spectrum of antimicrobial activity against different microorganisms including bacteria, viruses, fungi, and protozoa [1]. Human beta-defensins (β -defensins) are AMPs that not only protect mucosal membranes from microbial attack but also link two arms of immune system—innate and adaptive immunity—by performing numerous immune-related functions including activation, multiplication, and movement of immune cells to the area of infection, balancing of cytokine production, and regulation of repair and wound healing mechanisms [2]. Thus, defensins play a pivotal role in sustaining physiological homeostasis and act as a key guard of the tissues of the oral cavity, respiratory tract, skin, and digestive and genitourinary tracts against the pathogens [3]. Defensin's antimicrobial action is dependent on their cationic charge and also amphipathic organization of structure that enables it to bind to negatively charged bacterial membranes and then permeabilizing them to form pores that destroy bacteria [4]. One of the most important defensins is β -defensin 1 (*DEFB1* gene), and polymorphisms in the *DEFB1* gene have been involved and related with several diseases including allergies, cystic fibrosis, and cancer [5]. Multiple SNPs have been identified in the *DEFB1* gene, and such SNPs may modify the expression or activity of β -defensin 1 and could affect the susceptibility of individuals to infection [6]. In this study, we systematically collected missense nsSNPs of the *DEFB1* gene and screened the nsSNPs using multiple bioinformatics software tools to analyze the damaging nature of the nsSNPs of the *DEFB1* gene.

Methods

Dataset collection

The NCBI SNP database (<https://www.ncbi.nlm.nih.gov/SNP/>) was used to access the SNPs of the *DEFB1* gene (accessed 16 May 2020). The primary sequence of the protein (Uniprot accession number: P60022) encoded by the *DEFB1* gene was obtained from the UniProt database (accessed 16 May 2020). Only missense nsSNPs were chosen from the NCBI SNP database as they can modify the sequence of the amino acid encoded by the protein and have the potential to disturb the structural arrangement and function of the proteins.

Prediction of deleterious nsSNPs by various bioinformatics tools

Several online-based tools were employed to determine deleterious missense nsSNPs associated with the *DEFB1* gene. First, missense nsSNPs of the *DEFB1* gene selected from the NCBI SNP database were subjected to Sorting Intolerant from Tolerant (SIFT; <http://sift.bii.a-star.edu.sg/>) and Polymorphism Phenotyping v2 (PolyPhen v2; <http://genetics.bwh.harvard.edu/pph2/>) tools. SIFT makes a

prediction whether substitution of an amino acid has damaging effect on the function of the protein based on sequence alignment and homology [7]. SIFT gives a probability score of observing a new amino acid at that specific position, and a score less than or equal to the threshold of 0.05 is deleterious, and a prediction greater than the threshold is tolerant [7–9]. PolyPhen v2 predicts the possible impact of nsSNPs based on sequence, structural conformation, and phylogenetic features characterizing the substitution, and the output of the PolyPhen v2 is a prediction showing the substitution as probably damaging, possibly damaging, or benign, along with a numerical score ranging from 0.0 (benign) to 1.0 (damaging) [10].

The nsSNPs identified to be damaging by both SIFT and PolyPhen v2 tools were then subjected to Protein Variation Effect Analyzer (PROVEAN; <http://provean.jcvi.org>), Screening for non-acceptable polymorphisms (SNAP; <https://roslab.org/services/snap/>), Predictor of human deleterious single nucleotide polymorphism (PhD-SNP; <http://snps.biofold.org/phd-snp/phd-snp.html>), and SNPs&GO (GO-Gene Ontology; <http://snps-and-go.biocomp.unibo.it/snps-and-go/>) web-based tools. PROVEAN analyzes the consequence of sequence variation on the function carried out by the protein [11], and PROVEAN score ≤ -2.5 indicates that the amino acid variant has damaging effect whereas variant having score > -2.5 is considered to have neutral effect on the protein. SNAP is a neural network-based method, and using in silico-acquired sequence-based information classifies all nsSNPs into damaging and neutral [12]. PhD-SNP is based on support vector machines (SVMs) which predicts whether a point mutation is a neutral polymorphism or is associated to genetic disorders in humans [13]. SNPs&GO can determine whether a variation is disease-associated or not by gathering unique information derived from sequence of the protein, phylogenetic relation, and function encoded by the protein [14].

Prediction of nsSNPs effect on protein stability by I-Mutant 3.0

I-Mutant 3.0 tool was used to determine the impact of nsSNPs on protein stability (<http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>). I-Mutant 3.0 is trained to predict the effect of SNPs on the free energy change value (Delta Delta G: DDG) based on the sequence or tertiary structure of the protein [13, 15]. Delta Delta G (DDG) is the change in Gibbs free energy, and the change in the free energy of folding can be predicted from the difference in the free energies of the folded wild-type and mutant structures [16]. I-Mutant 3.0 gives the output as a DDG value which is categorized into one of the three predictions: largely unstable (DDG < -0.5 kcal/mol), largely stable (DDG > 0.5 kcal/mol), or neutral ($-0.5 \leq \text{DDG} \leq 0.5$ kcal/mol).

Phylogenetic conservation analysis of nsSNPs by ConSurf

The extent of evolutionary rate of conservation of amino acids in β -defensin 1 was analyzed using the ConSurf web server (<http://consurf.tau.ac.il/>). ConSurf calculates conservation scores of the amino acid position of the protein based on the phylogenetic relation between homologous sequences [17, 18]. ConSurf predicts conservation scores ranging from 1 to 9 with grade 9 being the most conserved position.

Post-translational modification prediction by ModPred

ModPred server (<http://www.modpred.org/>) predicts different types of post-translational modification (PTM) sites based on sequence [19]. PTM sites in β -defensin 1 protein sequence were predicted by the ModPred server.

Gene-gene interaction

Following the identification of several disease-associated polymorphisms by whole-genome association analysis, there is an increasing interest in the detection of the effects of polymorphism due to interaction with other genetic factors [20]. GeneMANIA (<http://www.genemania.org>) predicts gene function and generates information such as gene co-expression, co-localization, shared protein domains, and pathway involved [21]. The gene-gene interaction network of the *DEFB1* gene was predicted by GeneMANIA.

Prediction of protein structure and mutant analysis by the Project HOPE software tool

Have (y)Our Protein Explained (HOPE) is a web-based approach for the automated analysis of mutants and analyzes the consequence of point mutations on the structural conformation and function of proteins [22]. HOPE gathers information from various sources including annotated sequence data from the UniProt database and builds an extensive report about the effect of mutation with text, figures, and animations.

Results

Retrieval of nsSNPs from the NCBI SNP database

The nsSNPs of the *DEFB1* gene systematically examined in this study were retrieved from the NCBI SNP database. A total of 4024 SNPs were reported in the human *DEFB1* gene in the database, and among the SNPs reported, 86 were missense SNPs, 32 were synonymous, 45 SNPs in 5' untranslated region (UTR) region, 49 SNPs in 3' UTR, 2866 were intronic, and the rest were upstream and downstream transcripts. The missense nsSNPs were selected for our investigation since deleterious nsSNPs could have structural and functional impact on the protein.

Prediction of deleterious nsSNPs

First, 86 missense SNPs were submitted to the SIFT software tool which predicted that 18 nsSNPs were deleterious with score ≤ 0.05 and remaining were tolerant. Next, nsSNPs were subjected to PolyPhen v2 analysis. The results of both SIFT and PolyPhen v2 tools were combined to increase the accuracy of prediction, and only SNPs with SIFT score ≤ 0.05 and PolyPhen score > 0.90 were selected. Ten nsSNPs were identified by both tools as deleterious. The selected nsSNPs were submitted to other bioinformatics tools—PROVEAN, SNAP, PhD-SNP, and SNPs&GO. The results obtained from these tools are shown in Table 1. Eight nsSNPs were determined to be deleterious by all in silico tools mentioned above: rs1800968, rs140503947, rs145468425, rs146603349, rs199581284, rs201260899, rs371897938, rs376876621. rs55874920 was found to be deleterious by SIFT, PolyPhen v2, PROVEAN, and SNAP. rs56270143 was found to be deleterious by SIFT, PolyPhen v2, PROVEAN, SNAP, and PhD-SNP. These SNPs potentially predicted as deleterious were taken into consideration and further analyzed by other tools.

Table 1 List of nsSNPs of the *DEFB1* gene predicted as deleterious by different bioinformatics tools

SNP ID	Amino acid change	SIFT score	SIFT prediction	PolyPhen score	PolyPhen prediction	PROVEAN score	PROVEAN prediction	SNAP prediction	PhD-SNP prediction	SNPs&GO prediction
rs1800968	C67S	0	Deleterious	0.999	Probably damaging	- 9.697	Deleterious	Effect	Disease	Disease
rs55874920	T58S	0	Deleterious	0.967	Probably damaging	- 3.879	Deleterious	Effect	Neutral	Neutral
rs56270143	G62W	0	Deleterious	0.999	Probably damaging	- 6.121	Deleterious	Effect	Disease	Neutral
rs140503947	Y35C	0	Deleterious	0.984	Probably damaging	- 8.636	Deleterious	Effect	Disease	Disease
rs145468425	G57R	0	Deleterious	1	Probably damaging	- 7.758	Deleterious	Effect	Disease	Disease
rs146603349	P50Q	0	Deleterious	0.904	Possibly damaging	- 8.00	Deleterious	Effect	Disease	Disease
rs199581284	C44R	0	Deleterious	0.999	Probably damaging	- 12.00	Deleterious	Effect	Disease	Disease
rs201260899	C44Y	0	Deleterious	0.999	Probably damaging	- 11.00	Deleterious	Effect	Disease	Disease
rs371897938	G21R	0	Deleterious	0.948	Probably damaging	- 4.242	Deleterious	Effect	Disease	Disease
rs376876621	C59Y	0	Deleterious	0.999	Probably damaging	- 10.667	Deleterious	Effect	Disease	Disease

Prediction of effects of nsSNPs on protein stabilization by I-Mutant 3.0

I-Mutant server was used to analyze whether the selected missense nsSNPs increase or decrease the stability of β -defensin 1 protein. According to I-Mutant server, nsSNPs rs140503947 and rs146603349 showed a DDG value of less than -0.5 , which indicated that they were largely unstable and decreased protein stability. The results of I-Mutant server are shown in Table 2.

Evolutionary conservation analysis by ConSurf

Specific positions of amino acids that are essential for the correct function of a protein change more slowly than other residues (i.e., are evolutionarily conserved), and deleterious mutations are more frequent in conserved positions [23]. Selected nsSNPs analyzed with ConSurf revealed that they were positioned in the conserved regions, and the tool predicted that C67S, G57R, P50Q, C44R, C44Y, and C59Y were highly conserved. ConSurf also indicated that the selected nsSNPs have structural and functional effects on β -defensin 1 protein. C67S, T58S, and P50Q were predicted as functional residues, and G57R, C44R, C44Y, and C59Y were predicted as structural residues. The results are shown in Table 3.

Prediction of post-translational modification by ModPred

PTMs play a significant role in protein folding and degradation, regulation of gene expression, and in several biological pathways. An extensive analysis between SNPs, PTMs, and diseases are necessary as SNPs may induce PTMs, and determining disease-related nsSNPs associated with PTM sites can help in better interpretation of diseases [24]. PTMs mediated by nsSNP were analyzed using ModPred server. According to ModPred, disulfide linkage sites were predicted at C67S, C44R, and C59Y residues, and the results are shown in Table 3.

Table 2 DDG values as predicted by I-Mutant 3.0

SNP ID	Amino acid change	DDG value
rs1800968	C67S	-0.42
rs55874920	T58S	-0.48
rs56270143	G62W	-0.24
rs140503947	Y35C	-1.13
rs145468425	G57R	-0.44
rs146603349	P50Q	-1.12
rs199581284	C44R	-0.49
rs201260899	C44Y	-0.32
rs371897938	G21R	-0.24
rs376876621	C59Y	0.07

Gene-gene interaction

It has become highly important to predict genes with specific DNA sequence polymorphisms, each with combinations of wild-type and variant alleles and genotypes that impact susceptibility to a disease mainly through interactions with genetic and environmental factors [25]. GeneMANIA constructs a composite gene-gene functional interaction network, and the interaction network of the *DEFB1* gene as predicted by GeneMANIA is shown in Fig. 1.

Prediction of the effects of amino acid changes on β -defensin 1 protein from Project Hope

The Project HOPE software showed how the wild-type and variant amino acids contradicted from each other in terms of their physicochemical properties such as hydrophobicity, charge, and size. Predictions regarding the differences in the properties of the amino acids and the effect of polymorphic amino acids on the domain and conservation obtained from the Project HOPE software are given in Table 4. Results of Project HOPE for the 3D predictions of β -defensin 1 protein (wild and variant) are shown in Table 5.

Discussion

SNPs can act as markers in pharmacokinetics predicting responsiveness to treatment therapies and play a role in disease prognosis. It can also help in facilitating more tailored personalized treatment to patients and improve medication strategies. By studying the effects generated by functional coding SNPs in disease-associated proteins, new compounds can be designed for correcting the effects of those mutations in the population [26]. Several in silico studies on polymorphisms to predict nsSNPs as deleterious or neutral have been carried out. Studies have been carried out on different genes such as tumor necrosis factor- α (TNF- α) [27], ATP-binding cassette sub-family B member 1 (ABCB1) [28], B cell lymphoma/leukemia 11A (BCL11A) [29], and interleukin 27 [30].

In this study, various computational approaches were used to predict the deleterious nsSNPs of β -defensin 1. First, this work tested the missense nsSNPs of the *DEFB1* gene using SIFT and PolyPhen. The nsSNPs which were considered as deleterious by both these tools were then analyzed using PROVEAN, SNAP, PHD-SNP, and SNP&GO. By comparing the scores of these methods, ten nsSNPs were found to be deleterious—rs1800968 (C67S), rs55874920 (T58S), rs56270143 (G62W), rs140503947 (Y35C), rs145468425 (G57R), rs146603349 (P50Q), rs199581284 (C44R), rs201260899 (C44Y), rs371897938 (G21R), rs376876621 (C59Y). These ten SNPs were further analyzed by I-Mutant server which determined that nsSNPs rs140503947 and rs146603349 decreased protein stability. ConSurf analysis revealed that SNPs were located in the conserved regions. Defensins are amphipathic with a net positive charge and possess disulfide bonds which

Table 3 Prediction of phylogenetic conservation and post-translational modifications

SNP ID	Amino acid change	ConSurf conservation score	Functional/structural	Post-translational modification
rs1800968	C67S	9/conserved	f	Disulfide linkage
rs55874920	T58S	8/conserved	f	Amidation, O-linked glycosylation
rs56270143	G62W	6/intermediately conserved	-	-
rs140503947	Y35C	7/conserved	-	Amidation
rs145468425	G57R	9/conserved	s	-
rs146603349	P50Q	9/conserved	s	-
rs199581284	C44R	9/conserved	s	Disulfide linkage
rs201260899	C44Y	9/conserved	s	-
rs371897938	G21R	7/conserved	-	-
rs376876621	C59Y	9/conserved	s	Disulfide linkage

According to ConSurf server, "f" is functional residue and "s" is structural residue

are important for their molecular function [31]. ModPred results showed that SNPs affect PTMs such as disulfide linkage, amidation, and glycosylation. When disulfide linkages are affected, the native and misfolded defensins could exhibit altered biological functions affecting its antimicrobial and chemotactic activity [32]. The damaged PTMs by SNPs may affect protein's structure and functions and could be considered as biomarker candidates and drug targets for therapeutic purpose [33]. Gene-gene interaction network of the *DEFB1* gene predicted by

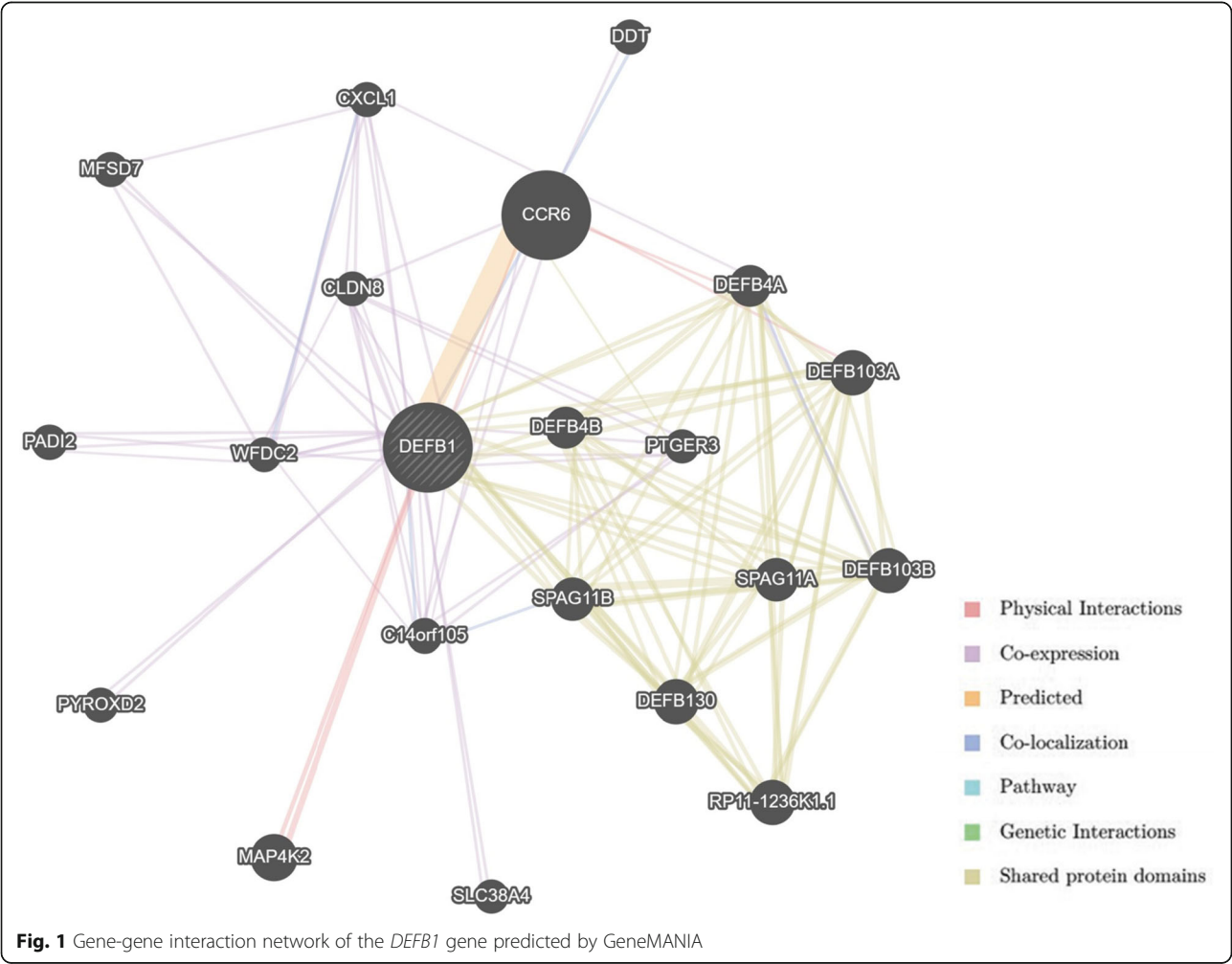


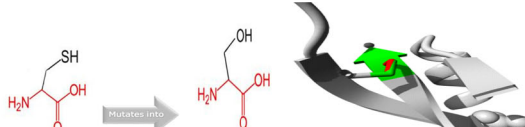
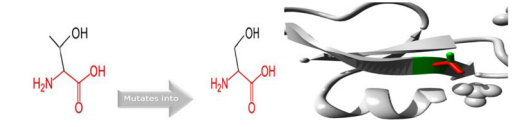
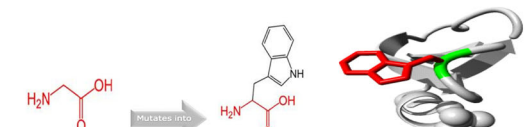
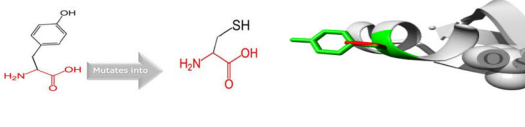
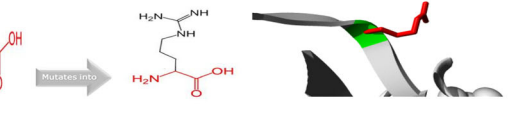
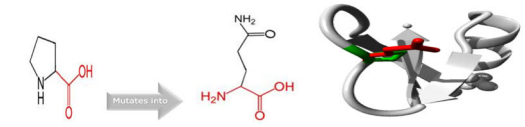
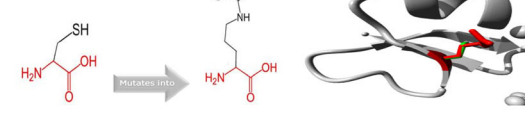
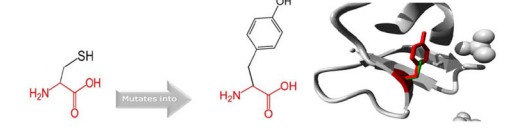
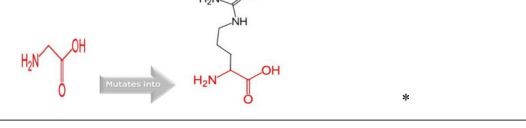
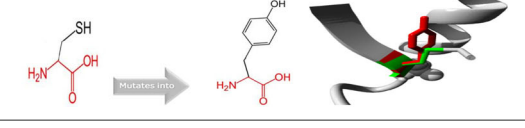
Table 4 Predictions based on the Project HOPE tool

SNP ID	Amino acid change	Amino acid properties	Location	Effect of polymorphism on the protein
rs1800968	C67S	Variant residue is less hydrophobic compared to wild-type residue; this polymorphism might disrupt hydrophobic associations in the core structural organization of the protein.	Wild-type is buried, and the variant might interrupt the core organization of this domain.	Location of the variant residue is near a conserved position and so this variation is likely to be deleterious.
rs55874920	T58S	Variant is smaller compared to wild-type amino acid and will lead to a loss of surface reactions.	Variant amino acid is positioned on the surface of a domain with undefined function.	This polymorphism is likely not damaging as homologous proteins exist with similar residues as the variant at this position.
rs56270143	G62W	Wild-type amino acid is smaller than the variant. As it is also positioned on the surface of the protein; variant can alter its external relations with other residues.	Variant amino acid is positioned on the surface of a domain with undefined function.	Wild-type amino acid is conserved, and based on conservation scores, this variant is likely deleterious.
rs140503947	Y35C	Wild-type amino acid is bigger, and the smaller variant will lead to possible loss of surface interactions.	Variant amino acid is positioned on the surface of a domain with undefined function.	Other residues have been noticed at this position. So, this variant might be found without being deleterious to the protein.
rs145468425	G57R	Wild-type amino acid is neutral in charge, and the variant is positive. This can cause repulsive reactions between the variant and neighboring residues.	Variant is positioned on the surface of a domain with undetermined role.	The amino acid is located near a conserved position, and this variant is likely damaging.
rs146603349	P50Q	Wild-type amino acid is smaller compared to the variant and is situated on the protein's surface. The variant can interrupt interactions with other residues.	Proline is considered to be rigid and produces a specific backbone conformation. This variation has the potential to alter this distinct conformation.	Closely related residues as the variant are found at this location in other similar proteins, and this polymorphism is probably not damaging.
rs199581284	C44R	As the variant is charged, it can affect folding of the protein. In the interior of the protein, hydrophobic interactions are disrupted as the wild-type amino acid is hydrophobic than the variant.	The amino acid is buried, and the variant possibly alters the core organization of this domain.	This variation is likely damaging.
rs201260899	C44Y	Variant is bigger than the wild-type and might not fit in the core of the protein.	The amino acid is buried, and the variant might disrupt the core structure of the domain.	This variation is likely damaging.
rs371897938	G21R	As the variant is a charged residue, this can give rise to repulsion problems with charged ligands or other residues.	New residue is introduced in the signal peptide and may disturb recognition of this signal peptide. Glycine is flexible compared to other residues; this flexible nature might be important for the protein's activity. Variant can destroy this function.	Other residues have been seen at this position. So, this variation might happen without being deleterious to the protein.
rs376876621	C59Y	Variant might not fit in the core as it is bigger than wild-type. Hydrophobic interactions are disturbed in the interior as the wild-type residue is more hydrophobic than the variant.	Wild-type is buried, and the variant might alter the core organization of this domain.	Only this residue type is found at this position, and this variation is likely damaging.

GeneMANIA showed that it interacts with CCR6, a chemokine receptor expressed by the immune cells such as dendritic cells and memory T cells. β -defensins recruit these immune cells to the site of microbial invasion by

interacting with CCR6 [34]. The *DEFB1* gene is constitutively expressed by epithelial cells of many tissues and so the depletion of its activity can increase the risk of microbial infection. Another study speculated that the

Table 5 Results of the 3D predictions of β -defensin 1 protein by Project HOPE. The protein is colored gray; the side chains of both the wild-type and the mutant residue are shown and colored green and red respectively

1. rs1800968 (C67S)	2. rs55874920(T58S)
	
3. rs56270143 (G62W)	4. rs140503947 (Y35C)
	
5. rs145468425 (G57R)	6. rs146603349 (P50Q)
	
7. rs199581284 (C44R)	8. rs201260899 (C44Y)
	
9. rs371897938 (G21R)	10. rs376876621 (C59Y)
	

*Due to a lack of structural information, HOPE could not generate 3D prediction for G21R polymorphism

downregulation of the *DEFB1* gene in various tumors such as renal cell carcinoma, prostate cancer, and nonfunctioning pituitary adenoma is mediated by the activation of phosphatidylinositol 3-kinase (PI3K)/protein kinase B (Akt)/mammalian target of rapamycin (mTOR) (PI3K-AKT-mTOR) pathway through transcriptional regulation of the *DEFB1* gene [35]. PI3K signaling pathway alters *PAX2* expression which in turn affects the *DEFB1* gene expression as *PAX2* acts as a transcriptional repressor of the *DEFB1* gene [35, 36]. Deleterious SNPs of the *DEFB1* gene may affect the interaction and functioning of other genes in the gene-gene interaction network. Physicochemical and molecular properties of the variant residues are sometimes different from the original structure, and the conservation of the structure is important for the preservation and execution of specific functions [37]. This encouraged us to study how the molecular properties differ from wild-type and variant amino acids using Project HOPE analysis. Project HOPE determined the possible effects of nsSNPs on the various molecular properties of the amino acid residues including size, charge, and hydrophobicity and also

predicted the impact of SNPs on the domain and structure of β -defensin 1 protein. The prediction of deleterious SNPs has been done through computational tools, and well-designed experimental and clinical analyses are required to examine the effects of these nsSNPs on structure and function of β -defensin 1 protein.

Conclusion

Several web-based algorithmic tools based on sequence and structural conservation were used to identify deleterious nsSNPs of the *DEFB1* gene. Using different computational tools, 86 missense nsSNPs of the *DEFB1* gene were screened; out of these, ten nsSNPs were found to be potentially deleterious. The deleterious variants in the *DEFB1* gene may affect its cellular function in protecting against microbial attack and also its immunomodulatory function. Use of multiple bioinformatics tools to predict pathogenic nsSNPs would be advantageous in reducing cost and time but confirmation of the roles of these nsSNPs requires experimental validation. This in silico study could form the basis for targeting pathogenic sites of β -defensin 1 protein.

Abbreviations

ABCB1: ATP-binding cassette sub-family B member 1; AMPs: Antimicrobial peptides; BCL11A: B cell lymphoma/leukemia 11A; DDG: Delta Delta G; *DEFB1*: Beta defensin 1; GO: Gene Ontology; HOPE: Have (y)Our Protein Explained; mTOR: Mammalian target of rapamycin; NCBI: National Center for Biotechnology Information; nsSNP: Non-synonymous single nucleotide polymorphism; PhD-SNP: Predictor of human deleterious single nucleotide polymorphism; PI3K: Phosphatidylinositol 3-kinase; PolyPhen: Polymorphism Phenotyping; PROVEAN: Protein Variation Effect Analyzer; PTM: Post-translational modification; SIFT: Sorting Intolerant from Tolerant; SNAP: Screening for non-acceptable polymorphisms; SNP: Single nucleotide polymorphism; TNF: Tumor necrosis factor; UTR: Untranslated region

Acknowledgements

We wish to thank the DST-FIST (Ref. No.SR/FST/College-23/2017), Government of India, New Delhi, India, for providing the funded research equipment facilities to Sree Balaji Dental College and Hospital, Pallikaranai, Chennai, Tamil Nadu, India. We also thank Sree Balaji Dental College and Hospital, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India for supporting the study.

Authors' contributions

Conception and design: US, PRB, and HVS. Literature search: HVS and US. Manuscript preparation: HVS. Editing and review: US and PRB. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public or commercial sectors.

Availability of data and materials

All data analyzed during this study are included in this article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Authors declare no conflict of interest.

Author details

¹Human Genetics Research Centre, Sree Balaji Dental College and Hospital, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India. ²Department of Genetic Engineering, Bharath Institute of Science and Technology, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India.

Received: 11 September 2020 Accepted: 3 November 2020

Published online: 29 November 2020

References

- Heimlich DR, Harrison A, Mason KM (2014) Host antimicrobial peptides in bacterial homeostasis and pathogenesis of disease. *Antibiotics*. 3(4):645–676
- Ghosh SK, McCormick TS, Weinberg A (2019) Human beta defensins and cancer: contradictions and common ground. *Front Oncol*. 9:341
- Meade KG, O'Farrelly C (2019) β -defensins: farming the microbiome for homeostasis and health. *Front Immunol*. 9:3072
- Dale BA, Krisanaprakornkit S (2001) Defensin antimicrobial peptides in the oral cavity. *J Oral Pathol Med Rev Artic*. 30(6):321–327
- de Oca EP-M (2010) Human β -defensin 1: a restless warrior against allergies, infections and cancer. *Int J Biochem Cell Biol*. 42(6):800–804
- Dale BA, Fredericks LP (2005) Antimicrobial peptides in the oral environment: expression and function in health and disease. *Curr Issues Mol Biol*. 7(2):119
- Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res*. 40(W1):W452–W457
- Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. *Genome Res*. 11(5):863–874
- Ng PC, Henikoff S (2002) Accounting for human polymorphisms predicted to affect protein function. *Genome Res*. 12(3):436–446
- Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 76(1):7–20
- Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012) Predicting the functional effect of amino acid substitutions and indels. *PLoS One*. 7(10):e46688
- Bromberg Y, Rost B (2007) SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res*. 35(11):3823–3835
- Capriotti E, Calabrese R, Casadio R (2006) Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics*. 22(22):2729–2734
- Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R (2009) Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat*. 30(8):1237–1244
- Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res*. 33(suppl_2):W306–W310
- Kellogg EH, Leaver-Fay A, Baker D (2011) Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins Struct Funct Bioinforma*. 79(3):830–838
- Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*. 38(suppl_2):W529–W533
- Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T et al (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*. 44(W1):W344–W350
- Pejaver V, Hsu W, Xin F, Dunker AK, Uversky VN, Radivojac P (2014) The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci*. 23(8):1077–1093
- Cordell HJ (2009) Detecting gene–gene interactions that underlie human diseases. *Nat Rev Genet*. 10(6):392–404
- Wardle-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P et al (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*. 38(suppl_2):W214–W220
- Venselaar H, Te Beek TA, Kuipers RKP, Hekkelman ML, Vriend G (2010) Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinf*. 11(1):1–10
- Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I et al (2013) ConSurf: using evolutionary data to raise testable hypotheses about protein function. *Isr J Chem*. 53(3–4):199–206
- Kim Y, Kang C, Min B, Yi G-S (2015) Detection and analysis of disease-associated single nucleotide polymorphism influencing post-translational modification. *BMC Med Genomics*. 8(S2):S7
- Gilbert-Diamond D, Moore JH (2011) Analysis of gene-gene interactions. *Curr Protoc Hum Genet*. 70(1):1–14
- Javed R (2010) Current research status, databases and application of single nucleotide polymorphism. *Pak J Biol Sci*. 13(13):657–663
- Dabhi B, Mistry KN (2014) In silico analysis of single nucleotide polymorphism (SNP) in human TNF- α gene. *Meta gene*. 2:586–595
- Chakraborty R, Gupta H, Rahman R, Hasija Y (2018) In silico analysis of nsSNPs in ABCB1 gene affecting breast cancer associated protein P-glycoprotein (P-gp). *Comput Biol Chem*. 77:430–441
- Das SS, Chakravorty N (2020) Identification of deleterious SNPs and their effects on BCL11A, the master regulator of fetal hemoglobin expression. *Genomics*. 112(1):397–403
- Koohyanizadeh F, Karaji AG, Falahi S, Rezaeimanesh A, Salari F (2020) In silico prediction of deleterious single nucleotide polymorphisms in human interleukin 27 (IL-27) gene. *Meta Gene*. Published online. 25:100710
- Park MS, Kim JI, Lee I, Park S, Bae JY, Park MS (2018) Towards the application of human defensins as antivirals. *Biomolecules & therapeutics*. 26(3):242
- Wang G (2012) Post-translational modifications of natural antimicrobial peptides and strategies for peptide engineering. *Current biotechnology*. 1(1):72–79
- Sun B, Zhang M, Cui P, Li H, Jia J, Li Y, Xie L (2015) Nonsynonymous single-nucleotide variations on some posttranslational modifications of human

proteins and the association with diseases. *Comput Math Methods Med.* 1: 2015

34. Yang D, Chertov OB, Bykovskaia SN, Chen Q, Buffo MJ, Shogan J, Anderson M, Schröder JM, Wang JM, Howard OM, Oppenheim JJ (1999 Oct 15) β -defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science.* 286(5439):525–528
35. Lee M, Wiedemann T, Gross C, Leinhäuser I, Roncaroli F, Braren R, Pellegata NS (2015) Targeting PI3K/mTOR signaling displays potent antitumor efficacy against nonfunctioning pituitary adenomas. *Clinical Cancer Research.* 21(14): 3204–3215
36. Bose SK, Gibson W, Bullard RS, Donald CD (2009) PAX2 oncogene negatively regulates the expression of the host defense peptide human beta defensin-1 in prostate cancer. *Mol Immunol.* 46(6):1140–1148
37. Castro HC, Sathler P (2015) Defensins and bioinformatics: in silico approaches for novel therapeutic antimicrobial peptides. *Elys J Microbes* 1(1):102

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)