INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print), 2222-5234 (Online)
http://www.innspub.net
Vol. 19, No. 1, p. 85-96, 2021

RESEARCH PAPER OPEN ACCESS

Prediction and evaluation of Deleterious Non-Synonymous SNPs (nsSNPs) in human IL17A Gene

Awad Ahmed Algarni

Department of Biology, College of Science and Arts, Albaha University, Baljurashi, Saudi Arabia

Key words: IL17A, Cancer, Mutations, nsSNPs, Polymorphisms.

http://dx.doi.org/10.12692/ijb/19.1.85-96 Article published on July 31, 2021

Abstract

Single nucleotide polymorphisms in the IL17A gene are associated with many types of cancer. Therefore, the identification of functional and structural polymorphisms in IL17A is important to study and determine therapeutic targets and their potential impact. In this study, several computational methods have been used to identify non-synonymous SNPs that are deleterious in the IL17A gene, including SIFT, PolyPhen2, PROVEAN, SNAP. PhD-SNP, SNP&GO and Pmut were used to predict disease-associated nsSNPs. I-mutant and Mupro were used to predict protein stability. ConSurf was used to predict conserved residues and functional regions of the protein. Finally, SOPMA followed by Project Hope software was used to predict the effect of these mutations on protein structure and function. Our study concludes that ten nsSNPs (R69Q, A92S, C94S, R95C, C99F, D107G, V121I, V142M, T145I and P149L) were shown to be potentially deleterious in IL17A. The present study represents a comprehensive in silico analysis of the IL17A gene and will be a useful tool for future studies.

* **Corresponding Author:** Awad Ahmed Algarni ✉ awad-algarni-j@hotmail.com

Introduction

Interleukin-17A (IL-17A) is one of a six-member family of pro-inflammatory cytokines (IL17A, IL17B, IL17C, IL17D, IL17E and IL17F) (Kolls and Lindén, 2004). The Interleukin-17A gene consists of 3 exons and 2 introns and is situated on human chromosome number 6p12 (Duan *et al.*, 2014; Lv *et al.*, 2015). Numerous studies demonstrated that polymorphism of the IL17A gene is linked to ulcerative colitis and Crohn's disease (Kim *et al.*, 2011; Zhang *et al.*, 2013). In addition, numerous recent studies have also shown the association of the IL17A polymorphism with susceptibility to ulcerative colitis (Arisawa *et al.*, 2008), rheumatoid arthritis (Nordang *et al.*, 2009), and different cancer including bladder cancer (Zhou *et al.*, 2013), breast cancer (Wang *et al.*, 2012), lung cancer (Cheng *et al.*, 2015), colorectal cancers (Omrane *et al.*, 2014), and gastric cancer (Meng *et al.*, 2010). Genetic diversity has a significant impact on gene function and is a major indicator of disease risk, including cancer. Single nucleotide polymorphism (SNP) is the most frequent type of human genome variation, resulting from the substitution of a single base pair (Barnes, 2010). Approximately 1.4 million SNPs are found in the human genome (Whiffin *et al.*, 2019). SNPs are a potential biomarker in predicting many diseases, such as cancer, Alzheimer's disease (Sirisena *et al.*, 2019). Many computational approaches are performed to predict missense variants, which may lead to amino acid substitutions that can affect protein stability (Thusberg *et al.*, 2011). The human gene IL17A's deleterious nsSNPs have not been predicted in silicon analysis. Therefore, the current study aimed to identify the pathogenic SNPs in IL17A and evaluate their impact on the structural and functional level using various computational approaches.

Materials and methods

Datasets

IL17A gene Data were obtained from the NCBI database and the IL17A protein amino acid sequence (accession: Q16552) was obtained from the UniProt database (Sherry *et al.*, 2001). The present study was used missense nsSNPs only.

SIFT

SIFT is a homology-based sequencing tool. SIFT can tell whether the function of a protein is influenced by the substitution of amino acids. Moreover, SIFT's high ability to discriminate between neutral and deleterious nsSNPs made it an important prediction tool (Kumar, 2009). All nsSNPs of the IL17A gene collected from dbSNP were submitted in SIFT tool to determine the deleterious nsSNPs (Ng and Henikoff, 2003).

Polyphen2

Polyphen2 is a web server that predicts the effect of amino acid substitutes on protein structure and function using physical and comparative considerations. The input for Polyphen2 was the protein sequence of the gene IL17A and amino acid variants. Based on the PSIC value "position-specific independent count" Polyphen2 classified nsSNPs in three classes: benign, probably damaging and possibly damaging (Adzhubei *et al.*, 2010).

PROVEAN

PROVEAN is a software tool that predicts the impact of amino acid substitution on protein function using alignment-based output. If the score is below 2.5, the variation will be deleterious, whereas the variant should be neutral if the score is above 2.5 (Choi *et al.*, 2012).

SNAP-2

SNAP-2 is a network method for predicting functional changes in SNPs based on structural and biophysical properties of amino acids in proteins (Bromberg *et al.*, 2009).

PhD-SNP

PHD-SNP is software that serves to predict whether a protein mutation at a particular point can be considered a disease or a neutral polymorphism (Capriotti *et al.*, 2006).

SNP&GO

SNP&GO is an algorithmic tool. It can predict if a variant is associated with a disease by using the

corresponding protein functional annotation (Calabrese *et al.*, 2009).

P-MuT

P-mut is software that predicts the potentially pathogenic nature of a nonsense mutation by a neural network method (Ferrer-Costa *et al.*, 2005).

ConSurf

ConSurf is a software tool that calculates the conservation of positions of amino acids depend on phylogenetic relationships among homologous sequences (Ashkenazy *et al.*, 2010).

HOPE

HOPE is a web-based tool that collects information from various sources such as the UniProt database. it is used to study the structural impact of the substitution of amino acids (Venselaar *et al.*, 2010).

Results and discussion

SNP Dataset and SNP retrieval

The IL17A gene Polymorphism data were obtained from the dbSNP database. A total of 1805 SNPs were found in the human IL17A gene. Among them, there are 112 missense SNPs, 64 were synonymous, 302 in 3′ UTR, 18 in 5′ UTR, 551 were intronic (Fig. 1). For further analysis, Missense SNPs have been selected.

Functional deleterious SNPs of IL17A

Among the 112 missense nsSNPs, SIFT analysis predicted that 42 (37.5%) nsSNPs were deleterious and 70 (62.5%) were tolerated. PolyPhene-2 software tool predicted 15 (13.4%) as probably damaging, 26 (23.2%) as possibly damaging, and 71 (63.4%) as benign mutations. PROVEAN software predicted that 46 (41%) nsSNPs as deleterious and 66 (59%) as neutral. SNAP-2 predicted that 81(72%) nsSNPs ware affect the normal function of proteins whereases 31 (28%) proved to be neutral. Out of 112 missense SNPs examined using SIFT, Polyphene-2, PROVEAN and SNAP2, it was indicated that 29 mutations were deleterious to IL17A protein.

This 29 nsSNPs were considered and analyzed further with other tools (Table. 1).

Table 1. SNPs analyzed by SIFT, Polyphen 2.0, PROVEAN and SNAP2 servers.

		SIFT		Polyphen		PROVEAN		SNAP	
SNP ID	Variants	prediction	score	prediction	score	prediction	score	prediction	score
rs761536762	T2A	DAMAGING	0.00	benign	0.000	Neutral	-0.596	neutral	-42
rs1457371293	P3A	TOLERATED	0.72	benign	0.016	Neutral	-0.806	effect	13
rs1179111332	G4R	TOLERATED	0.71	benign	0.003	Neutral	-0.356	neutral	-8
rs1410140371	K5N	TOLERATED	0.05	benign	0.025	Neutral	-0.073	effect	8
rs1421825821	T6P	DAMAGING	0.03	benign	0.002	Neutral	-1.151	effect	40
rs1163865638	T6I	DAMAGING	0.02	benign	0.002	Neutral	-0.932	neutral	-11
rs1328913681	L8S	TOLERATED	0.35	benign	0.001	Neutral	-0.777	effect	53
rs1394860452	V9A	TOLERATED	0.54	benign	0.001	Neutral	-1.030	effect	19
rs762765742	L11P	DAMAGING	0.00	possibly damaging	0.871	Deleterious	-2.930	effect	63
rs750881577	L13Q	DAMAGING	0.00	probably damaging	0.957	Deleterious	-3.466	effect	34
rs1255951047	L15V	TOLERATED	0.06	probably damaging	0.991	Neutral	-1.337	neutral	-55
rs1236360666	S16N	TOLERATED	0.20	benign	0.004	Neutral	-1.032	effect	4
rs751616243	S16R	TOLERATED	0.19	benign	0.154	Neutral	-1.326	effect	14
rs1191322787	A19T	TOLERATED	0.16	benign	0.055	Neutral	-1.572	neutral	-33
rs369300495	I20T	TOLERATED	0.29	benign	0.001	Neutral	-0.293	neutral	-14
rs1348882769	I27V	TOLERATED	0.69	benign	0.001	Neutral	0.160	neutral	-40
rs1187930738	I27T	TOLERATED	0.39	benign	0.003	Neutral	-0.055	neutral	-11
rs139620979	R29G	DAMAGING	0.02	benign	0.002	Neutral	-2.148	effect	27
rs144233360	R29Q	TOLERATED	1.00	benign	0.001	Neutral	0.190	neutral	-45
rs757736597	N30Y	DAMAGING	0.03	benign	0.240	Deleterious	-2.794	effect	33
rs1285825975	G32E	TOLERATED	1.00	benign	0.018	Neutral	-2.323	effect	45
rs746503257	S36P	TOLERATED	0.21	benign	0.001	Neutral	-1.191	neutral	-35
rs768149285	D38V	TOLERATED	0.07	benign	0.012	Deleterious	-3.918	effect	32

rs776322859	K39E	TOLERATED	0.27	benign	0.038	Neutral	-0.556	effect	18
rs1365509391	K39R	TOLERATED	0.24	benign	0.002	Neutral	-0.649	neutral	-12
rs747773286	N40T	TOLERATED	0.07	benign	0.122	Neutral	-1.375	neutral	-6
rs769613834	F41Y	TOLERATED	0.07	possibly damaging	0.767	Neutral	-1.367	effect	62
rs762757093	P42L	TOLERATED	1.00	benign	0.017	Neutral	-2.050	neutral	-5
rs566045823	R43W	DAMAGING	0.01	benign	0.332	Neutral	-1.641	effect	56
rs780218416	R43Q	TOLERATED	0.94	benign	0.000	Neutral	0.796	neutral	-29
rs759761787	T44P	TOLERATED	0.17	benign	0.201	Neutral	-2.219	effect	41
rs767695440	T44I	TOLERATED	0.10	benign	0.054	Neutral	-2.058	effect	7
rs997763562	N50S	TOLERATED	0.89	benign	0.013	Neutral	-1.031	neutral	-23
rs1327295035	H52R	TOLERATED	0.35	benign	0.000	Neutral	1.008	effect	44
rs754178712	R54W	DAMAGING	0.04	benign	0.332	Neutral	-2.099	effect	47
rs201890924	R54Q	TOLERATED	0.50	benign	0.001	Neutral	-0.304	effect	28
rs1245219600	T56N	TOLERATED	0.52	benign	0.070	Neutral	0.122	effect	38
rs1435810177	T58P	TOLERATED	0.26	benign	0.005	Neutral	-0.556	effect	19
rs953491556	N59Y	TOLERATED	0.05	benign	0.006	Neutral	-1.015	effect	2
rs746361396	N59T	TOLERATED	0.32	benign	0.080	Neutral	-0.002	neutral	-43
	N59S	TOLERATED	1.00	benign	0.004	Neutral	0.229	neutral	-36
	N59I	TOLERATED	0.05	benign	0.003	Neutral	-0.872	effect	6
rs1195936847	K61R	TOLERATED	0.84	benign	0.002	Neutral	-0.160	neutral	-24
rs754531736	D65H	DAMAGING	0.01	probably damaging	0.909	Deleterious	-3.543	effect	42
rs780646538	D65G	DAMAGING	0.04	benign	0.038	Deleterious	-4.073	effect	53
rs148704956	Y66C	DAMAGING	0.04	probably damaging	0.960	Deleterious	-3.569	effect	52
rs769455698	Y67H	TOLERATED	0.55	benign	0.001	Neutral	0.943	effect	4
rs948797711	N68S	TOLERATED	0.34	benign	0.056	Deleterious	-3.347	neutral	-19
rs777429640	N68K	TOLERATED	1.00	benign	0.005	Deleterious	-3.310	neutral	-24
rs747814163	R69Q	DAMAGING	0.00	probably damaging	0.999	Deleterious	-3.622	effect	75
rs1213956454	T71P	DAMAGING	0.04	probably damaging	0.995	Deleterious	-3.412	effect	39
rs138238811	T71I	TOLERATED	0.27	probably damaging	0.987	Neutral	-0.830	neutral	-28
rs578005242	W74S	DAMAGING	0.00	probably damaging	1.000	Deleterious	-13.428	effect	83
rs1177256325	N75K	TOLERATED	0.93	benign	0.123	Neutral	-1.946	effect	7
rs1380289273	L76I	TOLERATED	0.16	possibly damaging	0.666	Neutral	-0.500	effect	52
rs1401456522	L76P	DAMAGING	0.02	probably damaging	0.964	Deleterious	-3.361	effect	73
rs777451627	R78C	DAMAGING	0.03	possibly damaging	0.877	Neutral	-0.672	effect	42
rs760902102	R78H	DAMAGING	0.01	possibly damaging	0.828	Neutral	-1.067	effect	44
	R78P	DAMAGING	0.03	benign	0.019	Neutral	-0.850	effect	53
	R78L	TOLERATED	0.41	benign	0.033	Neutral	1.403	effect	16
rs768767102	N79S	TOLERATED	0.33	possibly damaging	0.589	Neutral	-1.436	neutral	-7
rs1217461322	Y85D	DAMAGING	0.00	probably damaging	0.995	Deleterious	-6.267	effect	80
rs372751013	S87A	TOLERATED	0.15	benign	0.021	Neutral	-1.584	neutral	-13
rs765332494	W90R	TOLERATED	0.13	possibly damaging	0.876	Deleterious	-3.033	effect	67
rs151317528	E91K	TOLERATED	0.05	possibly damaging	0.879	Deleterious	-3.267	effect	61
rs1180318768	E91D	TOLERATED	0.11	benign	0.289	Neutral	-1.967	effect	33
rs763243057	A92S	DAMAGING	0.00	probably damaging	1.000	Deleterious	-2.933	effect	51
rs1259842626	C94S	DAMAGING	0.00	probably damaging	0.99	Deleterious	-9.600	effect	79
rs199987410	R95C	DAMAGING	0.04	probably damaging	0.971	Deleterious	-5.124	effect	56
rs755641354	R95H	TOLERATED	0.12	probably damaging	0.937	Deleterious	-3.254	effect	62
	R95L	TOLERATED	0.40	benign	0.065	Deleterious	-3.550	effect	28
rs371175650	L97V	TOLERATED	0.20	benign	0.043	Neutral	-0.624	neutral	-16
rs753474484	G98S	TOLERATED	0.08	benign	0.218	Deleterious	-4.450	effect	55
rs868361380	G98A	TOLERATED	0.11	benign	0.132	Deleterious	-4.433	effect	53
rs1168004594	C99F	DAMAGING	0.00	probably damaging	1.000	Deleterious	-10.567	effect	87
rs375068948	I100L	TOLERATED	0.19	benign	0.007	Neutral	-1.033	effect	4
rs1474496126	N101Y	DAMAGING	0.01	possibly damaging	0.811	Deleterious	-6.143	effect	57
rs988912244	N101I	DAMAGING	0.00	possibly damaging	0.895	Deleterious	-6.531	effect	53
rs756965539	A102T	DAMAGING	0.02	benign	0.020	Deleterious	-2.522	effect	28
	A102P	TOLERATED	0.10	benign	0.026	Deleterious	-3.027	neutral	-3
rs778660554	N105K	TOLERATED	1.00	benign	0.002	Neutral	2.001	neutral	-62

rs745674559	V106M	TOLERATED	0.18	possibly damaging	0.845	Neutral	-0.459	neutral	-35
rs1347847159	D107G	DAMAGING	0.03	possibly damaging	0.716	Deleterious	-5.781	effect	72
rs780196923	H109R	TOLERATED	0.46	possibly damaging	0.856	Neutral	-1.965	effect	25
rs932255819	M110R	DAMAGING	0.00	benign	0.012	Deleterious	-3.572	effect	62
rs139375510	P114A	TOLERATED	0.64	benign	0.004	Deleterious	-5.800	effect	53
rs768657689	P114R	TOLERATED	0.44	benign	0.281	Deleterious	-6.567	effect	64
	P114L	TOLERATED	0.45	benign	0.069	Deleterious	-7.444	effect	52
rs748124159	I115V	TOLERATED	0.26	benign	0.168	Neutral	-0.928	effect	7
rs770066194	I115N	DAMAGING	0.00	probably damaging	1.000	Deleterious	-6.644	effect	75
rs1452861628	I119V	TOLERATED	0.35	benign	0.323	Neutral	-0.594	effect	3
rs1386733908	V121I	DAMAGING	0.00	probably damaging	0.997	Neutral	-0.761	effect	53
rs766719721	R123S	DAMAGING	0.01	probably damaging	0.957	Deleterious	-4.785	effect	78
	R123C	DAMAGING	0.00	probably damaging	0.986	Deleterious	-7.009	effect	66
rs145530358	R123H	DAMAGING	0.01	probably damaging	0.956	Deleterious	-4.186	effect	63
rs1235568790	R124K	TOLERATED	0.22	possibly damaging	0.869	Neutral	-2.367	effect	21
rs1336318401	E125K	TOLERATED	0.68	possibly damaging	0.783	Neutral	-1.949	effect	6
rs767872393	P126T	TOLERATED	0.59	benign	0.054	Neutral	-2.354	neutral	-47
rs1343283331	H128R	TOLERATED	0.63	benign	0.193	Neutral	-0.439	effect	47
rs756665674	P130R	DAMAGING	0.01	benign	0.045	Deleterious	-2.981	effect	10
rs1208310869	S132C	DAMAGING	0.02	probably damaging	0.937	Deleterious	-2.959	effect	27
	S132F	DAMAGING	0.01	benign	0.197	Deleterious	-3.170	effect	64
rs764987987	R134W	DAMAGING	0.00	probably damaging	0.970	Deleterious	-5.936	effect	53
rs147810050	R134Q	TOLERATED	0.23	benign	0.082	Neutral	-2.041	effect	17
rs199827182	L139P	DAMAGING	0.02	probably damaging	0.972	Deleterious	-4.640	effect	65
rs780929803	V142M	DAMAGING	0.00	probably damaging	0.998	Deleterious	-2.767	effect	76
rs1163763403	T145I	DAMAGING	0.00	probably damaging	1.000	Deleterious	-5.633	effect	63
rs1250966771	T148S	TOLERATED	0.42	benign	0.019	Neutral	-2.242	neutral	-11
rs749493493	T148N	TOLERATED	0.15	benign	0.040	Deleterious	-3.209	effect	21
rs868170851	P149L	DAMAGING	0.01	probably damaging	0.917	Deleterious	-8.289	effect	63
rs768084470	H152Q	TOLERATED	0.17	benign	0.000	Neutral	-1.074	neutral	-37
rs1308096563	A155S	TOLERATED	0.48	benign	0.002	Neutral	0.204	neutral	-18

Disease-associated nsSNPs of IL17A

The disease-related analysis included three prediction algorithms: PhD SNP, SNP&GO and P-mut. PhD-SNP analysis predicted that 20 nsSNPs are associated with disease and 9 nsSNPs are neutral. SNPs&GO analysis predicted that 23 nsSNPs are associated with disease and 6 nsSNPs are neutral. P-mut software predicted that 22 nsSNPs are associated with disease and 7 nsSNPs are neutral (Fig. 2).

Table 2. Protein stability analysis due to nsSNPs in human IL17A gene using I Mutant and MuPro.

	SNP ID	Variants	I Mutant			MuPro	
			Stability	RI	DDG (Kcal/mol)	Effect stability of protein structure	Confident Score
1	rs762765742	L11P	Increase	3	-0.90	DECREASE	-0.519
2	rs750881577	L13Q	Decrease	4	-1.27	DECREASE	-0.537
3	rs754531736	D65H	Decrease	6	-0.79	DECREASE	-1
4	rs148704956	Y66C	Decrease	3	-1.28	DECREASE	-1
5	rs747814163	R69Q	Decrease	9	-1.47	DECREASE	-0.273
6	rs1213956454	T71P	Decrease	4	-1.02	INCREASE	0.034
7	rs578005242	W74S	Decrease	9	-1.97	DECREASE	-0.771
8	rs1401456522	L76P	Decrease	3	-1.78	DECREASE	-1
9	rs777451627	R78C	Decrease	3	-1.10	DECREASE	-0.326
10	rs760902102	R78H	Decrease	9	-1.48	DECREASE	-0.600
11	rs1217461322	Y85D	Decrease	6	-1.24	DECREASE	-0.992

12	rs763243057	A92S	Decrease	7	-0.65	DECREASE	-1
13	rs1259842626	C94S	Decrease	7	-0.84	DECREASE	-0.569
14	rs199987410	R95C	Decrease	6	-0.90	DECREASE	-1
15	rs1168004594	C99F	Decrease	3	-0.04	DECREASE	-0073
16	rs1474496126	N101Y	Decrease	2	0.09	DECREASE	-0.073
17	rs988912244	N101I	Increase	2	0.89	DECREASE	-0.405
18	rs1347847159	D107G	Decrease	2	-1.22	DECREASE	-1
19	rs770066194	I115N	Decrease	6	-2.08	DECREASE	-0.935
20	rs1386733908	V121I	Decrease	1	-0.08	DECREASE	-1
21	rs766719721	R123S	Decrease	7	-1.04	DECREASE	-0.682
22		R123C	Decrease	4	-0.85	DECREASE	-0.475
23	rs145530358	R123H	Decrease	9	-1.24	DECREASE	-0.932
24	rs1208310869	S132C	Decrease	5	-0.77	DECREASE	-0.026
25	rs764987987	R134W	Decrease	6	-0.52	DECREASE	-0.335
26	rs199827182	L139P	Decrease	5	-1.62	DECREASE	-1
27	rs780929803	V142M	Decrease	9	-1.29	DECREASE	-0.308
28	rs1163763403	T145I	Increase	1	-0.12	DECREASE	-0.264
29	rs868170851	P149L	Decrease	3	-0.52	INCREASE	0.560

Prediction of protein stability

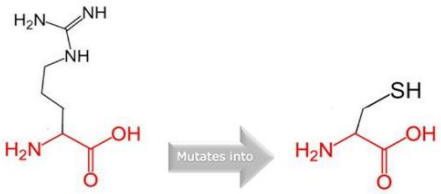
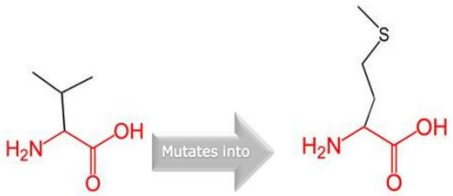
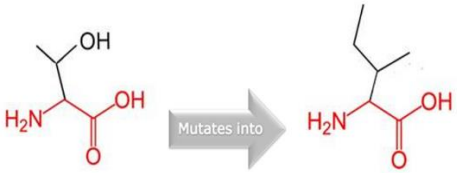
Table 2 shows that of the 29 nsSNPs, I-Mutant 2.0 had predicted 26 nsSNPs to result in decreased protein stability, while MUpro had predicted 27 nsSNPs to result in decreased protein stability.

Conservational analysis of IL17A

Amino acids involved in critical biological activities are conserved more than other residues. Therefore, nsSNPs are deleterious at highly preserved amino acid positions than nsSNPs at non-conversion sites (Doniger *et al.*, 2008). Ten variants of the 29 highly deleterious nsSNPs in the IL17A gene were found to be highly conserved. Four of them (R69Q, R95C, D107G, and P149L) are structural and buried, and six (A92S, C94S, C99F, V121I, V142M, and T145I) are functional and exposed with a scale score greater than 8 (Fig. 3a).

Table 3. Predictions based on the Project HOPE tool.

SNP ID	Amino acid change	Amino acid size	Charge	Change Hydrophobicity	Structural illustration of amino acid substitutions	Over impact
rs747814163	R69Q	W>M	W(positive) M (Neutral)	---		The mutant residue is smaller, this might lead to loss of interactions.
rs763243057	A92S	W<M	---	Decrease		Hydrophobic interactions, either in the core of the protein or on the surface, will be lost.
rs1259842626	C94S	---	---	Decrease		Hydrophobic interactions, either in the core of the protein or on the surface, will be lost.

rs199987410	R95C	W>M	W(positive) M (Neutral)	Increase		The mutant residue is smaller, this might lead to loss of interactions. The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding. The charge of the wild-type residue will be lost, this can cause loss of interactions with other molecules or residues.
rs1168004594	C99F	W<M	---	---		The mutant residue is bigger, this might lead to bumps
rs1347847159	D107G	W>M	W(negative) M (Neutral)	Increase		The mutant residue is smaller, this might lead to loss of interactions. The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding. The charge of the wild-type residue will be lost, this can cause loss of interactions with other molecules or residues.
rs1386733908	V121I	W<M	---	---		The mutant residue is bigger, this might lead to bumps
rs780929803	V142M	W<M	---	---		The mutant residue is bigger, this might lead to bumps
rs1163763403	T145I	W<M	---	Increase		The mutant residue is bigger, this might lead to bumps. The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding.
rs868170851	P149L	W<M	---	---		The mutant residue is bigger, this might lead to bumps

Secondary structure prediction

The secondary IL17A protein structure was predicted by SOPMA. The IL17A protein is consists of 196 amino acids (fig. 3b). Random coils were the highest number of sequence 79 (50.97%), 32 (20.65%) amino acids in the extended strand, 36 (23.23%) amino acids in the alpha helix, and 8 (5.16%) amino acids in the beta-turn(Fig. 3c). Among of 29 deleterious variants, fifteen variants (D65H, Y66C, R69Q, T71P, W74S, R78C, R78H, Y85D, C99F, N101Y, N101I, D107G, I115N , T145I and P149L) have been on the random coil, five variants (L11P, L13Q, A92S, C94S, and R95C) in the alpha-helix and nine variants (L76P, V121I, R123S, R123C, R123H, S132C, R134W, L139P and V142M) in the extended strand of the protein(Fig. 3b).

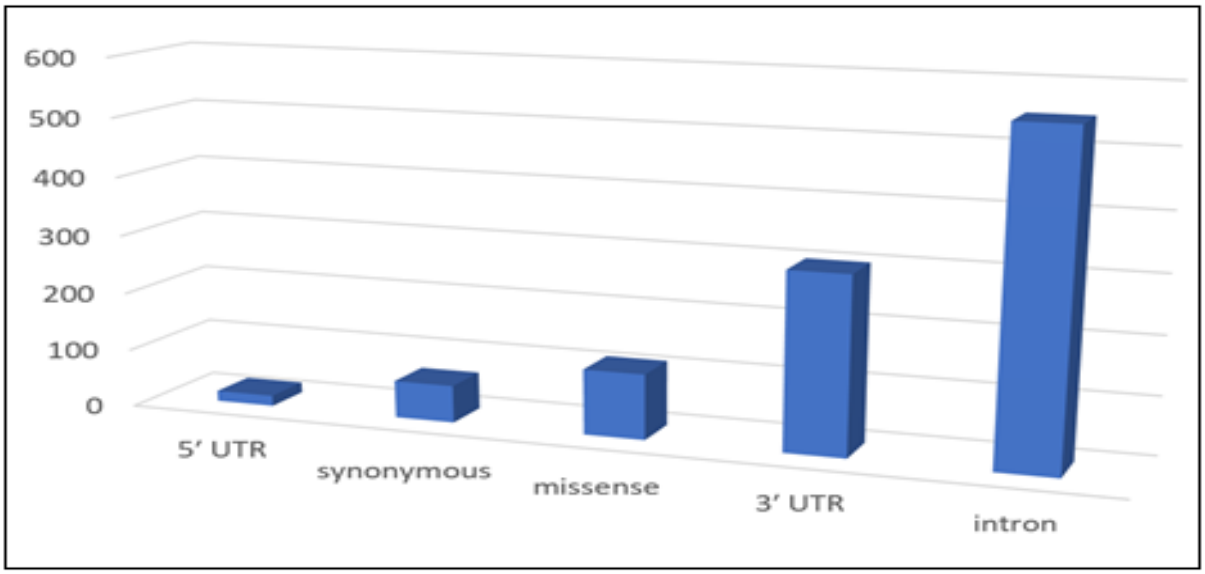


Fig. 1. Distribution of SNPs in different regions of IL17A gene.

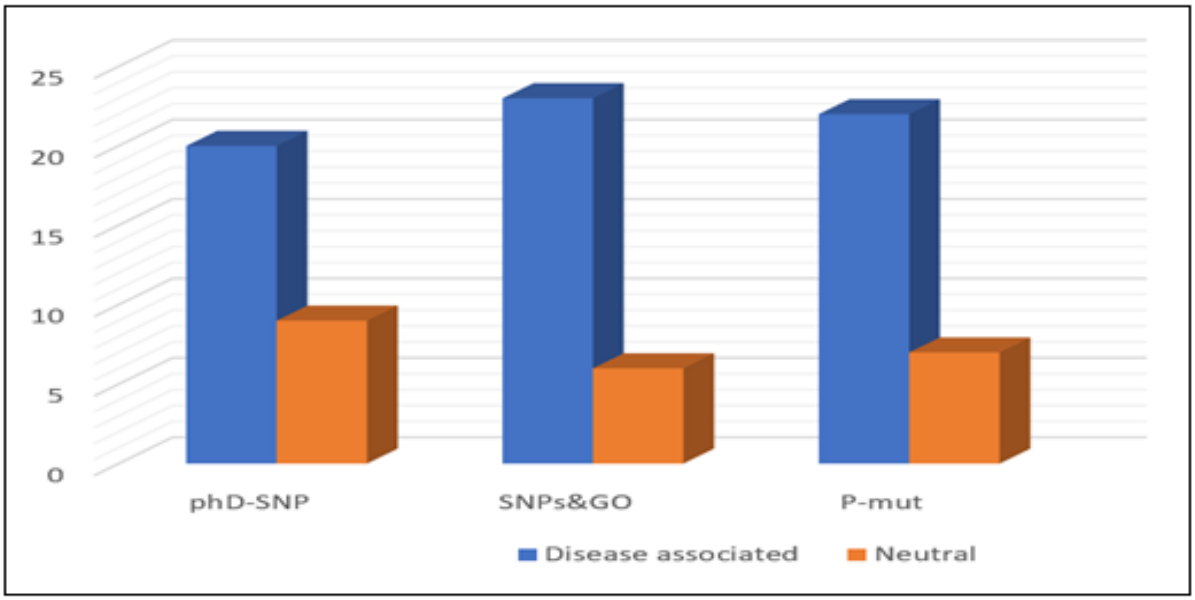


Fig. 2. Disease-associated nsSNPs predicted using PhD-SNP, SNPs & GO and P-mut tools.

Protein structure prediction

The Hope project was used to examine the effect of 10 functional and exposed SNPs with a score above 8 on the size, charge and hydrophobicity of amino acids. Three mutant amino acids were smaller in size than their wild-type counterparts, while five mutant amino acids were larger. Charge shifts occurred at three sites: two changed from positive to neutral, and one

changed from negative to neutral. TWO mutations Decrease hydrophobicity, and three increased hydrophobicity. This finding indicates that amino acid mutations at these locations affect protein function by changing the structure of the protein and modifying the protein's domain interactions with other molecules (Table 3).

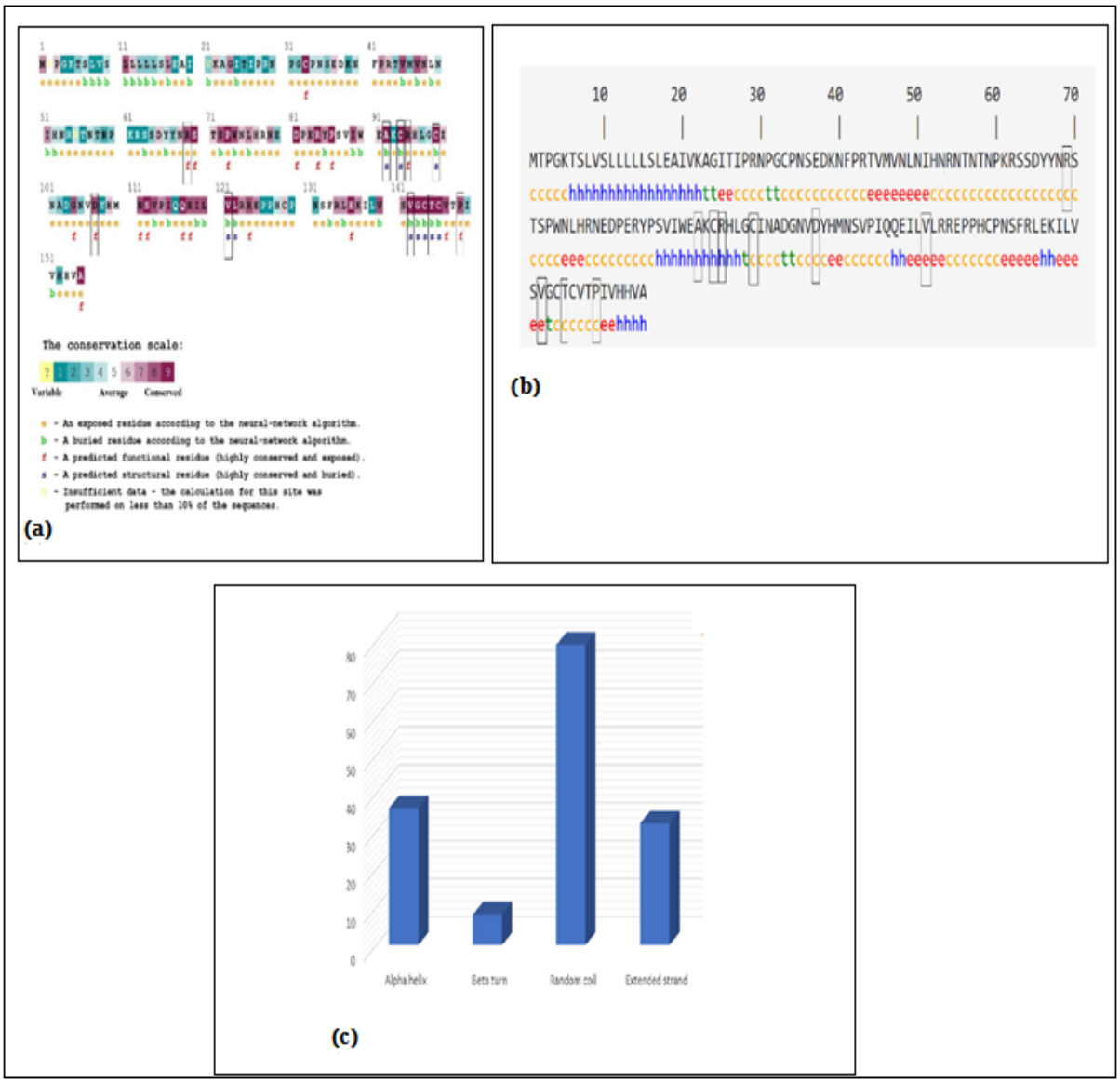


Fig. 3. (a) Analysis of evolutionarily conserved amino acid residues of by ConSurf. The color coding bar shows the conservation score, and boxes indicate the high-risk nsSNPs. (b) SOPMA analysis of the secondary structure of individual amino acid residues in protein produced from the IL17A gene. The highly deleterious nsSNPs are indicated through boxes. (c) SOPMA predicted the distribution of high-risk IL17A nsSNPs in random coils, alpha helixes, extended strands, and beta turns.

Conclusion

For the identification of deleterious IL17A nsSNPs, several sequences and structural software tools have been used. We analyzed 112 missense nsSNPs of the IL17 gene using various in silico methods; ten nsSNPs; rs747814163 (R69Q), rs763243057 (A92S), rs1259842626 (C94S), rs199987410 (R95C), rs1168004594 (C99F), rs1347847159 (D107G), rs1386733908 (V121I), rs780929803 (V142M), rs1163763403 (T145I) and rs868170851 (P149L) were shown to be potentially deleterious. Multiple bioinformatics tools would be beneficial for

predictions of pathogenic nsSNPs to reduce costs and times, while the roles of such nsSNPs need experimental confirmation.

Conflict of Interest

The author declares that there is no conflict of interest.

Funding

This research work has not received any funding at all in any part of the study.

Acknowledgments

The author would like to thank the Department of Biology, College of Science and Arts, Albaha University, Baljurashi, Saudi Arabia.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nature methods **7(4),** 248-249.
https://doi.org/10.1038/nmeth0410-248

Arisawa T, Tahara T, Shibata T, Nagasaka M, Nakamura M, Kamiya Y, Fujita H, Nakamura M, Yoshioka D, Arima Y, Okubo M. 2008. The influence of polymorphisms of interleukin-17A and interleukin-17F genes on the susceptibility to ulcerative colitis. Journal of clinical immunology **28(1),** 44-49.
https://doi.org/10.1007/s10875-007-9125-8

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 research **38,** W529-W533.
https://doi.org/10.1093/nar/gkq399

Barnes MR. 2010. Genetic variation analysis for biomedical researchers: a primer. Genetic Variation 1-20.
https://doi.org/10.1007/978-1-60327-367-1_1

Bromberg Y, Overton J, Vaisse C, Leibel RL, Rost B. 2009. In silico mutagenesis: a case study of the melanocortin 4 receptor. The FASEB Journal **23(9),** 3059-3069.
https://doi.org/10.1096/fj.08-127530

Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. 2009. Functional annotations improve the predictive score of human disease-related mutations in proteins. Human mutation **30(8),** 1237-1244.
https://doi.org/10.1002/humu.21047

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.
https://doi.org/10.1093/bioinformatics/btl423

Cheng S, Shao Z, Liu X, Guo L, Zhang X, Na Q, Chen X, Ma Y, Zheng J, Song B, Liu J. 2015. Interleukin 17A polymorphism elevates gene expression and is associated with increased risk of nonsmall cell lung cancer. DNA and cell biology **34(1),** 63-68.
https://doi.org/10.1089/dna.2014.2628

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS One* **7,** e46688.
https://doi.org/10.1371/journal.pone.0046688

Doniger SW, Kim HS, Swain D, Corcuera D, Williams M, Yang SP, Fay JC. 2008. A catalog of neutral and deleterious polymorphism in yeast. PLoS genetics **4(8),** p.e1000183.
https://doi.org/10.1371/journal.pgen.1000183

Duan Y, Shi JN, Pan C, Chen HL, Zhang SZ. 2014. Association between the interleukin-17A-197G> A (rs2275913) polymorphism and risk of digestive cancer. Asian Pacific Journal of Cancer Prevention **15(21),** 9295-9300.
https://doi.org/10.1007/s13277-014-1890-4

Ferrer-Costa C, Gelpí JL, Zamakola L, Parraga I, De La Cruz X, Orozco M. 2005. PMUT: a web-based tool for the annotation of pathological mutations on proteins. Bioinformatics **21(14),** 3176-3178.
https://doi.org/10.1093/bioinformatics/bti486

Kim SW, Kim ES, Moon CM, Park JJ, Kim TI, Kim WH, Cheon JH. 2011. Genetic polymorphisms of IL-23R and IL-17A and novel insights into their associations with inflammatory bowel disease. Gut **60(11),**1527-1536.
https://doi.org/10.1136/gut.2011.238477

Kolls JK, Lindén A. 2004. Interleukin-17 family members and inflammation. Immunity **21(14),** 467-476.
https://doi.org/10.1016/j.immuni.2004.08.018

Kumar A. 2009. An overview of nested genes in eukaryotic genomes. Eukaryotic Cell **8(9),** 1321-1329.
https://doi.org/10.1128/EC.00143-09

Lv Q, Zhu D, Zhang J, Yi Y, Yang S, Zhang W. 2015. Association between six genetic variants of IL-17A and IL-17F and cervical cancer risk: a case–control study. Tumor Biology **36(5),** 3979-3984.
https://doi.org/10.1007/s13277-015-3041-y

Meng XY, Zhou CH, Ma J, Jiang C, Ji P. 2012. Expression of interleukin-17 and its clinical significance in gastric cancer patients. Medical oncology **29(5),** 3024-3028.
https://doi.org/10.1007/s12032-012-0273-1

Ng PC, Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic acids research **31(13),** 3812-3814.
https://doi.org/10.1093/nar/gkg509

Nordang GB, Viken MK, Hollis-Moffatt JE, Merriman TR, Førre ØT, Helgetveit K, Kvien TK, Lie BA. 2009. Association analysis of the interleukin 17A gene in Caucasian rheumatoid arthritis patients from Norway and New Zealand. Rheumatology **48(4),** 367-370.
https://doi.org/10.1093/rheumatology/ken512

Omrane I, Marrakchi R, Baroudi O, Mezlini A, Ayari H, Medimegh I, Stambouli N, Kourda N, Bouzaienne H, Uhrhammer N, Bougatef K. 2014. Significant association between interleukin-17A polymorphism and colorectal cancer. Tumor Biology **35(7),** 6627-6632.
https://doi.org/10.1007/s13277-014-1890-4

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. Nucleic acids research **29(1),** 308-311.
https://doi.org/10.1093/nar/29.1.308

Sirisena ND, Samaranayake N, Dissanayake VH. 2019. Genotype data for single nucleotide polymorphism markers in sporadic breast cancer related genes in a Sri Lankan case–control cohort of postmenopausal women. BMC research notes **12(1),** 1-3.
https://doi.org/10.1186/s13104-019-4472-0

Thusberg J, Olatubosun A, Vihinen M. 2011. Performance of mutation pathogenicity prediction methods on missense variants. Human mutation **32(4),** 358-368.
https://doi.org/10.1002/humu.21445

Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. 2010. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC bioinformatics **11(1),** 1-10.
https://doi.org/10.1186/1471-2105-11-548

Wang L, Jiang Y, Zhang Y, Wang Y, Huang S, Wang Z, Tian B, Yang Y, Jiang W, Pang D. 2012. Association analysis of IL-17A and IL-17F polymorphisms in Chinese Han women with breast cancer. PloS one **7(3),** p. e34400.
https://doi.org/10.1371/journal.pone.0034400

Whiffin N, Ware JS, O'Donnell-Luria A. 2019. Improving the understanding of genetic variants in rare disease with large-scale reference populations. JAMA **322(13),** 1305-1306. https://doi.org/10.1001/jama.2019.12891

Zhang X, Yu P, Wang Y, Jiang W, Shen F, Wang Y, Tu H, Yang X, Shi R, Zhang H. 2013. Genetic polymorphisms of interleukin 17A and interleukin 17F and their association with inflammatory bowel disease in a Chinese Han population. Inflammation Research **62(8),** 743-750. https://doi.org/10.1007/s00011-013-0629-9

Zhou B, Zhang P, Wang Y, Shi S, Zhang K, Liao H, Zhang L. 2013. Interleukin-17 gene polymorphisms are associated with bladder cancer in a chinese han population. Molecular carcinogenesis **52(11),** 871-878. https://doi.org/10.1002/mc.21928