INT. J. BIOL. BIOTECH., 16 (4): 857-867, 2019.

STUDY OF TLR3 GENE NON-SYNONYMOUS SNP IN HUMAN: A COMPUTATIONAL APPROACH

Syeda Marium Ali[1], Khalida Naveed[1], Syed Noman Ali[2*], Syeda Hira Seemab[1], Mahwish Saleem[1] and Ayesha Majid[1]

[1]*Baqai Institute of Information Technology, Baqai Medical University, Karachi, Pakistan*
[2]*Department of Livestock, Government of Sindh, Pakistan*
Corresponding Author*: syednomanalivet@gmail.com

ABSTRACT

Toll like receptor is involved in detecting pathogen-associated molecular patterns of various pathogens and trigger an immune response against microbial infection. SNPs effect on structure and function of TLR3 gene was carried out using computational analysis during the current study. The 340 non synonymous Single nucleotide polymorphisms (nsSNPs) were evaluated either effecting function of TLR3 or using seven different tools employing specific threshold values. A total of four nsSNPs (L159F, L545P, N284I, and P880Q) were found to have deleterious effects on the TLR3 protein function and were therefore selected for further analysis. All of the variants showed increased free energy, although L159F exhibited the highest energy increase while the RMSD (root-mean-square deviation) values of L159F and L545P are greater than 0.15 strongly indicated structural changes. By using various publicly available computational tools, 5 deleterious mutations (L159F, L545P, N284I, L412F and P880Q) in the coding region of the TLR3 gene were identified. Structurally, the selected mutants (especially L159F, L545P and N284I) were predicted to be significant for pathogenicity. Among the 06 potential nsSNPs, the L159F, L545P and N284I mutants were found to cause decreased aggregation of the TLR3 protein. Therefore, these nsSNPs should be considered as important candidates in causing diseases related to TLR3 gene malfunction.

Key word: Bioinformatics**,** Single nucleotide polymorphism, Pattern Recognition Receptor,

Abbreviation: SNPs (Single nucleotide polymorphism)**,** nsSNPs (non-synonymous SNPs), PAMPs (pathogen-associated molecular patterns), TLRs (Toll-like receptors), RLRs (RIG-I-like receptors), NLRs (NOD-like receptors), CLRs (C-type lectin receptors), LRRs (leucine-rich repeats), TIR (Toll/interleukin-1 receptor), T1DM (Type 1 diabetes mellitus), IRF-3 (interferon regulatory factor 3), TLR3 (Toll-like receptor 3), RI (reliability index).

INTRODUCTION

The first line of defense includes innate immunity in against towards viruses, bacteria and fungi (Akira *et al*., 2006). This immunity provides essential cellular and molecular protection against invading pathogens (Tassia *et al*., 2017). The components of innate immunity are macrophages, dendritic cells, histiocytes, kupffer cells and mastocytes. These cells have a receptor called PRR (Pattern Recognition Receptor), which recognize the pattern on pathogen collectively referred to as PAMPs (pathogen-associated molecular patterns). Moreover, Gram-negative lipopolysaccharide or viral dsRNA, frequently serve basic biological roles in PAMPs (Tassia *et al*., 2017). The PRR include Toll-like receptors (TLRs), RIG-I-like receptors (RLRs), NOD-like receptors (NLRs), and C-type lectin receptors (CLRs) (Studzińska *et al*., 2017). TLR expression are found in cell types such as airway and gut epithelial cells, B cells, mast cells, NK cells, dendritic cells, regulatory T cells, macrophages, monocytes, neutrophils, basophils and endothelial cells (Pandey and Agrawal ., 2006). TLRs, was firstly originate as Toll protein in Drosophila melanogaster, these are a group of type-I transmembrane glycoprotein localized to plasma membranes and endosomes. Three main regions that possess in all TLR are: an extracellular domain of tandem leucine-rich repeats (LRRs), a transmembrane helix, and a cytoplasmic Toll/interleukin-1 receptor (TIR) domain (Botos *et al*., 2011). Human have ten different components from where (TLR1, 2, 4, 5, 6, and 10) recognize the host cell membranes and from (TLR3, 7, 8, and 9) recognize nucleic acids (Kawai and Akira, 2010). TLR3 expression has been found in endosomal compartments (such as dendritic cells and macrophages) or at the cell surfaces such as human fibroblasts (Bugge *et al*., 2017).

Toll-like receptor 3 (TLR3) initiate an innate immunity by recognizing pathogenic, double-stranded RNAs and stimulate downstream signaling through TIR domain containing adaptor TCAM1. It has a main role in antiviral defense and tumor eradication through the production of type I interferons (IFNs) and inflammatory cytokines (Salaun *et al*., 2006) (Salaun *et al*., 2007). TLR3 is involved in the activation of transcription factor nuclear factor

κB (NF- κB) and interferon regulatory factor 3 (IRF-3) that result in the production of inflammatory cytokines and interferon beta (IFNβ. Concurrent with TLR3 activation and interaction with TCAM1, another TIR domain containing protein, SARM1, increases in concentration and subsequently acts as a negative regulator for the TCAM1-dependent pathway (Yin and Gao, 2010). TLR3 gene is involved in the development of T1DM (Type 1 diabetes mellitus) and viral hepatitis infection and biliary cirrhosis (Assmann *et al*., 2015). Immunohistochemistry analyses revealed that the expression of TLR3 was markedly increased in biliary epithelial cells at sites of ductular reaction in primary biliary cirrhosis and autoimmune hepatitis (Nakamura *et al*., 2008). A strong positive correlation between the mRNA levels of TLR3 and type I IFN in the liver was found in the patients with primary biliary cirrhosis, suggesting TLR3 signaling is involved in the pathogenesis of primary biliary cirrhosis (Nakamura *et al*., 2008).

Basically, in the current study, we used different widely available bioinformatics tools and databases to explore nsSNPs in the TLR3 gene. Our aim is to find out the stable mutation in TLR3 gene and compared this structure with native protein to evaluate structural deviations and topological similarities.

MATERIALS AND METHODS

SINGLE NUCLEOTIDE POLYMORPHISM (SNP) Dataset: TLR3 protein sequence and their SNPs data were collected from protein and dbSNP-NCBI from the website "http:// www.ncbi.nlm.nih.gov/SNP/" hosted at NCBI. Additionally the protein sequence of the TLR3 gene was collected in FASTA format from the UniProtKB database website "http://www.uniprot.org/uniprot/".

SIFT TOOL: SIFT is a tool sequenced based for the comparison of amino acid sequence with the native amino acid. SIFT prediction is based on the degree of conserved amino acid residues whether the SNPs are deleterious or tolerated to characterize the effect of changes on protein structure and function. The SIFT program (version 2) is used to find out the amino acid substitution in TLR3 protein accessible at website "http://sift.jcvi.org/". The analyses of SIFT score was done on the basis of tolerance index (TI) score, for every SNP. A TI score _ 0.05 express the delirious effect of non-synonymous variants on protein function (Nakamura *et al*., 2008).

Screening for Non-acceptable Polymorphisms (SNAP2): SNAP2 server was used to predict the single amino acid substitutions in the TLR3 protein. SNAP2 server is a neural-network learning device from the website "https://rostlab.org/services/snap2web/". The result was noted in the form of (effect or neutral) and score were ranging from -100(strong neutral prediction) to +100(strong effect prediction). The reliability index (RI) score was also maintained for the accurate and less reliable mutation in the protein (input) (Kumar *et al*., 2009).

nsSNP Analyzer: nsSNP Analyzer is based on random forest, machine learning method. This server "http://snpanalyzer.uthsc.edu/" was used in the nsSNPs of the Tlr3 protein to predict the phenotypic effects whether it is associated with the disease or neutral. The input format of the sequence is in FASTA format (Bromberg *et al*., 2008).

Polymorphism Phenotyping, Version-2(PolyPhen-2): PolyPhen-2 server http://genetics.bwh.harvard.edu /pph2 / was used to predict possible impact of an amino acid substitution on the structural and functional of a TLR3 protein. The inputs were entered in FASTA sequence with detailed substitution of amino acid position. This server works on the basis of Naïve Bayes positive probability, report false positive and true positive result. The variants was evaluated on the basis of damaging and tolerated. The score from 0.85 to 1.0 was interpreted as damaging and from 0.15 to 1.0 was interpreted as possibly damaging and 0.0 to 0.15 was interpreted as tolerated (Bao *et al*., 2005).

FATHMM: The FATHMM tool (http://fathmm.biocompute.org.uk/inher- ited.html) was used predict the functional consequences of both coding variants and non-coding variants. This tool work on the basis of two algorithms that is weighted and un-weighted. Un-weighted worked on the basis of sequence/conservation based while weighted is worked on the basis of combined sequence conservation with pathogenicity weights. The weighted algorithm was used for further analysis of TLR3 protein.

PROVEAN (Protein Variation Effect Analyzer): The PROVEAN (http://provean.jcvi.org/index.php) tool was used to predict that amino acid substitution has an impact on the biological function of a protein. The input is in the FASTA format and SNP details. The value was evaluated on the basis of deleterious and mutual if the score is under

a threshold level than SNP is considered as deleterious and otherwise it will be predicted as neutral (Adzhubei *et al*., 2013).

Snpeffect: Snpeffect database 4.0 (http://snpeffect.switchlab.org) which is used to predict the Phenotyping effect on human SNPs, (Choi and Chan, 2015) basically the prediction purpose is based on the sequence- and structure-based tools such as TANGO to predict the aggregation-prone regions, (De Baets *et al*., 2011), WALTZ to predicts the amyloid-forming regions (Fernandez-Escamilla *et al*., 2004) , LIMBO that determines the chaperone binding sites for the Hsp70 chaperones (Maurer-Stroh *et al*., 2010) and finally, the FoldX server that was used to filter structurally destabilizing mutations of the wild-type protein (Van Durme *et al*., 2009). The result is based on PDB ID/Uniprotkb ID and SNP information.

I-Mutant 2.0: I-Mutant 2.0 (http://folding.biofold.org/cgi-bin/ i-mutant2.0.cgi), is an SVM-based server, which is used to analyzed the change in protein stability. The input limitation of this tool was given as the protein sequence, room temperature in Celsius degrees about [0-100] and pH value about [0-14]. It is predicted as the free energy change value (DDG) through which a point mutation stabilizes or destabilizes the native protein structure based on the free energy change (DDG) (Schymkowitz *et al*., 2005).

ConSurf: ConSurf server (http://consurf.tau.ac.il/2016/) is a bioinformatics web-based tool which identify the evolutionary conservation of amino acid positions in the TLR3 protein (Capriotti *et al*., 2005). ConSurf calculating the result of protein structure using empirical Bayesian interference or Maximum likelihood (ML) method. The input was in FASTA sequence on the basis of colored regions and divided into 9 conservation scale lengthen from variable, average and conserved. The amino acids with highly conserved scores are biologically predicted to be more significant.

MODELLER 9.17: Modeller 9.17 was used to predict the structural deviation and stability differences between native and mutant models of TLR3 protein. The crystal structure of the TLR3 protein was available in Protein Data Bank (PDB ID 1ZIW, and the resolution was about 2.40 Å) (Ashkenazy *et al*., 2010). This structure was used as a template for further analysis through MODELLER. The energy minimization, hydrogen bonding pattern and bond distances was initially analyzed by SWISS-PDB Viewer of modeled 3D structure variants using a version of the GROMOS 43B1 force field in the GROMOS96 software package integrated in the Swiss-PDB Viewer (Tanabe *et al*., 2015). TM-Align (http://zhanglab.ccmb.med.umich.edu/TM-align/) and the root mean square deviations (RMSDs) were used to calculate the scores (template modeling) (Johansson *et al*., 2012).

NetSurfP: The NetSurfP (http://www.cbs.dtu.dk/services/NetSurfP/) was used to identify the surface accessibility and secondary structure of amino acids in a protein. The input is in the FASTA-for- matted protein sequence. The result was examined according to solvent accessibility of the amino acids. (Zhang and Skolnick, 2005).

FTSite: Prediction of ligand binding sites was done by the server FTSite (http://ftsite.bu.edu/). It usually binds with small molecules of different shapes and polarity (Petersen *et al*., 2009). The input was entered as the menu having a job name, PDB ID (or file) and a protein chain, if a proteins contain multiple subunits.

String Database: String Database was used to identify the function of the interactions of a target protein with other corresponding proteins (Kozakov *et al*., 2015). Hence, this tool easily predicts both experimental and hypothetical interactions.

Identification of Post-translational Modification Sites on TLR3 protein: The Post-translational Modification identified on TLR3 protein by NetGlycate 1.0 server (http://www.cbs.dtu.dk/services/NetGly-cate/) and NetPhos2.0 server (http://www.cbs.dtu.dk/ services/NetPhos/). In NetGlycate 1.0 used to predict glycation sites on the e amino group of lysine residues in the TLR3 protein. (Szklarczyk *et al*., 2010) And Phosphorylation sites was identified in the primarily serine, Threonine and tyrosine residues of the TLR3 protein (Johansen, *et al*., 2006). In adding up, the Ubiquitylation sites of the protein were predicted using UbPerd (www.ubpred.org) (Blom *et al*., 1999). And SUMOylation sites in the TLR3 protein were predicted by SUMOplot tool, available at website (http://www.abgent.com/sumoplot).

RESULTS AND DISCUSSION

Retrieval of SNPs: SNPs of the TLR3 gene was retrieved from dbSNP (Database of single nucleotide polymorphisms (SNPs) in the exonic region of TLR3, 2568 SNPs were found and among them in which 168 were synonymous, 340 were non synonymous-missense, 13 were non synonymous-nonsense and 15 was a frame-shift mutation. In this computational analysis, non-synonymous-missense SNPs was chosen for further studies.

Prediction of Deleterious nsSNPs by SIFT, SNAP2 and nsSNPAnalyzer: The effect of nsSNPs on protein function was predicted by SIFT tool. According to this tool and prediction 15 out of 340 nsSNPs were filtered out of which 6 were damaging with low confidence (score of 0.00–0.02), while the remaining 09 variants were considered to be tolerated (score of 0.07–1.00). Among the 6 variants, the three nsSNPs (rs3775291), (rs76759876), (rs113258886) was predicted to be damaging with low confidence (score of 0.01) and (score of 0.02). In contrast, three nsSNPs (rs5743316, rs111488413, rs112666655) were predicted to be highly deleterious, with a TI score of 0.00. The 15 variants which was predicted from SIFT program were subjected to the SNAP2 and nsSNPAnalyzer. It showed 10 pathological polymorphisms were predicted 10 to be significant, while the remaining 5 nsSNPs were predicted to be neutral in SNAP2 server while in the nsSNPAnalyzer it showed 3 nsSNPs might be associated with disease causality while the 12 nsSNPs were found to be neutral. The overall deleterious nsSNPs was obtained from the SIFT, nsSNPAnalyzer and SNAP2 tools revealed that 4 nsSNPs (with rsIDs of rs76759876, rs112666655, rs5743316 and rs111488413 were commonly predicted to be deleterious. The results are shown in Table 1.

Table 1. Most suspicious nsSNPs found in the coding region of the TLR3 protein, identified using seven bioinformatics tools.

RS IDS	Snp IDS	SNAP2	POLYPHEN	SIFT	PROVEAN	Snp Analyzer	PhD-Snp	Fathmm
rs76759876	L159F	Effect	Probably damaging	Damaging	Deleterious	Neutral	Neutral	Tolerated
rs3775291	L412F	Neutral	Probably damaging	Damaging	Deleterious	Neutral	Disease	X
rs112666655	L545P	Effect	Probably damaging	Damaging	Deleterious	Disease	Disease	Tolerated
rs5743316	N284I	Effect	Probably damaging	Damaging	Deleterious	Disease	Disease	Tolerated
rs111488413	P880Q	Neutral	Probably damaging	Damaging	Deleterious	Disease	Disease	Tolerated
rs113258886	S258G	Effect	Benign	Damaging	Neutral	X	Neutral	X

Deleterious nsSNPs by PolyPhen-2 server: The fifteen nsSNPs obtained from the SIFT program were also submitted to the Polyphen-2 server. A total of 8 out of 15 variants were predicted as Probably Damaging with the PSID score with specificity and sensitivity (0.949-1.00) and 7 were predicted as Benign with the PSID score with specificity and sensitivity (0.001-0.269). The eight damaging variants in the Polyphen sever were matched with SIFT tool output and found that only 4 damaging variants were found both in the SIFT and Polyphen-2 server. These four variants were damaging which was predicted before by SIFT, nsSNP analyzer and SNAP2 tool.

Detection of deleterious nsSNPs by PhD-SNP, FATHMM and PROVEAN: The PhD-SNP, FATHMM and PROVEAN tools were used to confirm the deleterious and neutral SNP in the human TLR3 gene from the SIFT tool. The PhD-SNP server showed 4 out of 6 predicted to be damaging variants with the probability score greater than 0.5 and the remaining nsSNPs were to be neutral. From the PROVEAN tool, 5 nsSNPs (out of 6) were predicted to be deleterious, with the score below the threshold -2.5, which had scores above the threshold, were identified as neutral. In FATHMM tool it showed all the variants to be tolerated. So we combined the seven computational tools (SIFT, nsSNPAnalyzer, SNAP2, PolyPhen-2, PhD-SNP,FATHMM and PROVEAN), 6 nsSNPs

(L159F, L412F, L545P, N284I, P880Q, S258G) were found to be an related with the disease and from the 6 nsSNPs only 4 nsSNPs were used for further analysis.

Detection of Effective variants by Snpeffect: The effective mutant was detected through TANGO tool, WALTZ, LIMBO and FOLDX proteins and only WALTZ indicated that only 1 variant i.e. (N284I) had an increased amyloid forming propensity with the (dWALTZ score: 317.13), While the other tool did not show any effective mutation. The detailed results for the four selected variants are provided in Table 2.

Confirmation of Protein Stability by I-Mutant 2.0: The protein stability was also confirmed through the I-Mutant 2.0 server as shown in table 2. It was detected that 1 mutant (N284I) predict increased of protein stability with the RI index (0) and 3 mutants predicted decrease of the protein stability with the reliability index (7-9).

Table 2. Characterization of selected mutants by the Snpeffect and I-Mutant tools.

Amino acid change	SNP(Single Nucleotide Polymorphism) effect				I-Mutant	
	TANGO aggregation tendency (dTANGO score)	WALTZ amyloid propensity (dWALTZ score)	LIMBO chaperone binding tendency (dLIMBO score)	FoldX protein stability (ddG Kcal/mol)	Prediction	RI
L159F	NA (0.00)	NA (0.00)	NA (0.00)	Reduced (2.36)	Decrease	7
L545P	NA (0.00)	NA (0.00)	NA (-0.11)	Reduced (1.37)	Decrease	6
N284I	NA (36.68)	Increases (317.13)	NA (0.00)	Reduced (4.11)	Increase	0
P880Q	NA (2.38)	NA (0.24)	NA (0.00)	-------	Decrease	9

*NA-Not Affected

Function Prediction and Conservation Analysis by ConSurf:

The human TLR3 is highly conserved protein showed by ConSurf tool. Two protein mutants N284I and P880Q were located in highly conserved and were predicted to cause functional as well as exposed impact. While, the residue L159F and L545P was in conserved region and predicted to be buried (Fig. 1).

Structural Analysis of Mutant Structures:

The total energy of native and mutant model was calculated by SPDViewer. The total energy values for the 3-modeled mutant structures were shown in Table 3 and among three variants only L159F showed the increase in free energy, which may be due to substitution of (Phe) an aromatic amino acid from the (Leu) polar amino acid. Now further their TM-score and RMSD was calculated. The TM-score of the three variants were greater than 0.5 which revealed that there were no structural differences between the native and mutant modeled structures, while the RMSD values greater than 0.15 were considered significant, two mutants showed high RMSD values i.e. L159F (0.78) and L545P (0.64) which affect the function of protein and indicate the pathogenicity and these mutants shown in Table 4. The intermolecular forces such as H-bonding have an important effect on secondary structure of proteins (Radivojac, P. *et al*., 2010), seen by the Swiss-PDB Viewer 4.02. The mutant L159F showed a significant change in the bond distance when Leucine is replaced by Phenylalanine as compared to other mutant which did not show any significant difference in bond distance as shown in Table 5. Therefore it may be concluded that there is a decent chance of disruption in secondary structure of the L159F mutant. In adding up, the native and selected mutants of the TLR3 protein were analyzed for solvent accessibility and stability; significant changes in both parameters were strongly predicted for all 4 selected variants as shown in Table 6.

Fig.1. Unique and conserved regions in the TLR3 protein were determined using ConSurf. The color coding bar shows the color scheme representation of the conservation score.

Table 3. Total energy of native and mutant protein structures before and after energy minimization.

Variants	Total energy before energy minimization (kJ/mol)	Total energy after energy minimization (kJ/mol)
Native	-16999.215	-34746.188
L159F	-14087.378	-33757.109
L545P	-18161.199	-34914.063
N284I	-17894.434	-34825.730

Table 4. Structure alignment comparing mutant models and the wild-type 1ZIW model.

Variants	Tm-Align		
	Align	RMSD	TM-score
L159F	1ZIW	0.78	0.99213
L545P	1ZIW	0.64	0.99467
N284I	1ZIW	0.11	0.99984

Table 5. Surface accessibility of wild-type.

Amino Acid Change	Class assignment	Relative surface accessibility (RSA)	Absolute surface accessibility	Z-fit score
L159F	Buried	0.060	11.041	-1.300
	Buried	0.068	13.688	-1.484
L545P	Buried	0.246	45.006	-1.159
	Exposed	0.412	58.491	-1.351
N284I	Buried	0.083	12.166	-1.686
	Buried	0.089	16.446	-1.241

Table 6. Comparison of the hydrogen bonding pattern between the wild-type structure and selected mutants.

Position	Amino acid	Hydrogen bond contact	Bond distance
159	L	Leu159-Ile154 Leu159-Asn152 Leu159-Ile106	3.17 3.00 3.32
	F	Phe159- Ile154 Phe159-Asn152 Phe159- Ile106	3.14 3.30 3.35
545	L	LEU545-Ile531 LEU545-Asn533	3.40 2.98
	P	Pro545- Ile531 Pro545- Asn533	3.01 2.98
284	N	Asn284-Thr248 Asn284-Arg222	3.17 3.32
	I	Ile284- Thr248 Ile284- Arg222	3.15 3.34

Ligand Binding Sites and Protein–Protein Interactions:

FTSite was used to find the three ligand binding sites or active sites on the TLR3 protein given in Table 7 and it was observed that our selected variants were not involved in the ligand binding sites.

The STRING database was used to find out the protein-protein interaction and 10 intvaety patterns were interacted with the TLR3 protein shown in Fig. 2. There is no significant difference in the score bit the protein MYD88, TRAF6, TAB2, TICAM1, RIPK1,MAP3K7 and LYS96 showed that there is experimental evidence for the interaction of TLR3.

Table 7. Amino acid residues involved in ligand binding sites of the TLR3 protein.

Binding site 1	Binding site 2	Binding site 3
PHE A 227	PRO A 368	VAL A 625
ASN A 229	GLY A 369	PHE A 634
ASN A 252	ILE A 370	LEU A 640
SER A 254	SER A 392	MET A 642
SER A 256	LEU A 393	PHE A 647
MET A 278	ARG A 394	THR A 650
ASP A 280	THR A 395	CYS A 651
SER A 282	LEU A 396	GLU A 652
PHE A 304	THR A 397	SER A 653
GLU A 306	GLU A 399	VAL A 658
	THR A 400	TRP A 660
		THR A 664

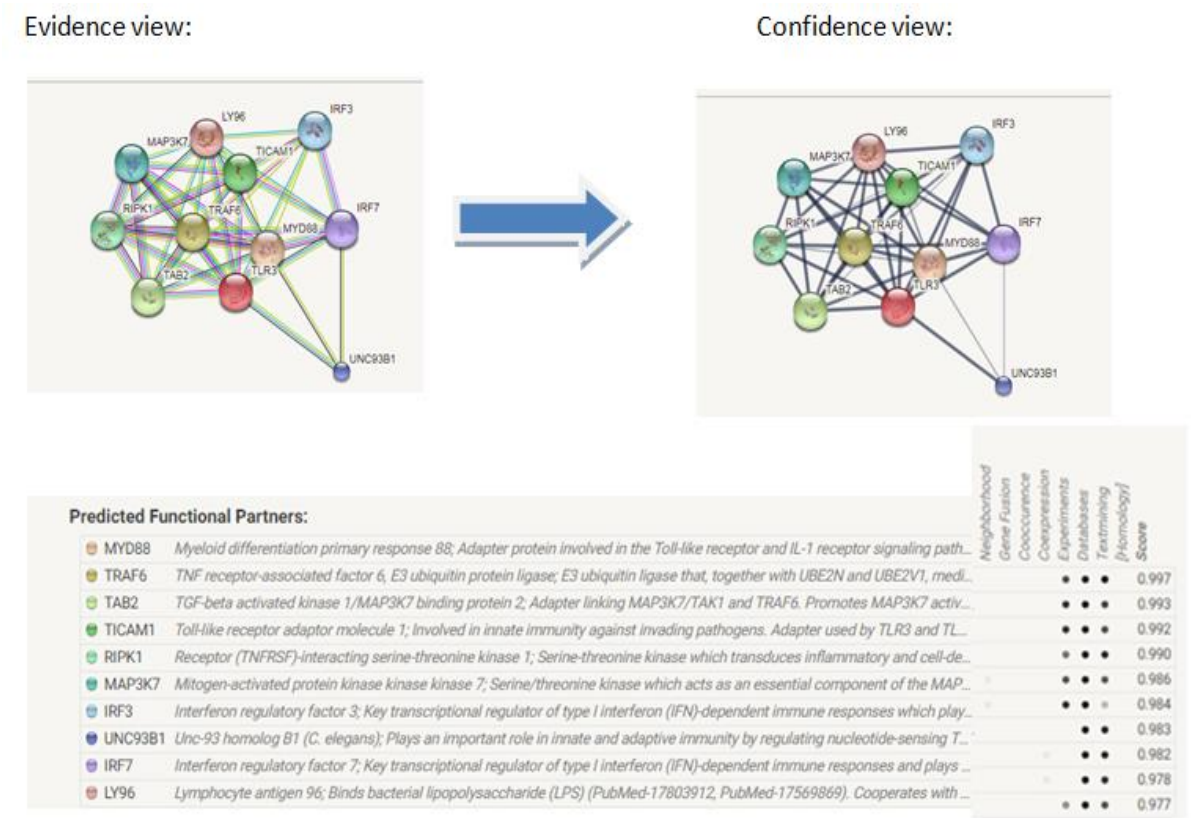

Fig.2. Protein-protein interactions of TLR3 with 10 partners. One color is assigned to each type of evidence predicted functional links.

Identification of Post-translational Modification Sites on the TLR3 Protein:

The nsSNPs were concluded by different in silico tools to determine whether any deleterious nsSNPs are present that would influence the post-translational modifications of the TLR3 protein. In this study, NetGlycate predicted that 19 residues undergo glycation. According to NetPhos, 32 serine, 7 Threonine and 5 tyrosine residues undergo Phosphorylation Table 8. The TLR3 protein also has 6 Lys residues in various positions in the sequence

were assigned a score above 0.00, indicating that these are possible Ubiquitylation sites. In SUMOylation sites, the SUMOplot was used in which we identified 9 different Lys residues Table 9. Similar to Ubiquitylation, Our selected 4 nsSNPs were not involved in the post-translational modification of the TLR3 protein.

Table 8. Post-translationally modified glycation and Phosphorylation sites in the TLR3 protein.

Glycation		**Phosphorylation**					
N-glycation		**Serine**		**Threonine**		**Tyrosine**	
Position	Score	Position	Score	Position	Score	Position	Score
41	0.784	29	0.985	4	0.755	356	0.564
89	0.910	48	0.997	25	0.799	375	0.826
145	0.828	80	0.991	43	0.966	517	0.642
147	0.717	164	0.889	74	0.709	808	0.548
187	0.904	204	0.528	127	0.728	907	0.783
201	0.961	210	0.544	135	0.559	-	-
210	0.892	237	0.951	287	0.515	-	-
272	0.562	247	0.922	-	-	-	-
335	0.912	262	0.668	-	-	-	-
418	0.961	285	0.610	-	-	-	-
493	0.978	298	0.705	-	-	-	-
613	0.896	303	0.973				
619	0.635	305	0.761	-	-	-	-
697	0.852	310	0.982	-	-	-	-
745	0.890	364	0.525				
779	0.680	388	0.526	-	-	-	-
785	0.835	434	0.657	-	-	-	-
808	0.816	436	0.787				
872	0.873	456	0.576				
	-	469	0.996	-	-	-	-
	-	477	0.849	-	-	-	-
	-	547	0.980				
		549	0.842				
-	-	660	0.565	-	-	-	-
-	-	673	0.941	-	-	-	-
-	-	743	0.696				
		744	0.986				
-	-	780	0.662	-	-	-	-
-	-	825	0.752	-	-	-	-
-	-	832	0.740				
-	-	855	0.971				
		859	0.778				

Table 9. Putative Ubiquitylation and SUMOylation sites.

Ubiquitylation			**SUMOylation**	
Score	**Position**	**Confidence**	**Lys position**	**Score**
163	0.66	Low	190	0.91
187	0.87	Medium	836	0.8
200	0.66	Low	350	0.61
201	0.68	Low	213	0.59
371	0.68	Low	539	0.5
697	0.62	Low	204	0.48
-	-	-	98	0.39
			118	0.15
			638	0.13

Conclusion

The 340 non-synonymous-missense region of TLR3 gene were passed through SIFT tool and find out 15 variants out of which 6 were damaging(L159F, L412F,N284I, P880Q, L545P, S258G). These damaging variants were confirmed by various bioinformatics tool and 4 variants (L159F, L545P, N284I and P880Q) were common in all tools and were predicted to be deleterious. Among these 4 variants N284I and P880Q showed amyloid propensity with the increase of protein stability and these 2 variants were also located in highly conserved as well as exposed region while the energy minimization showed us that L159F energy was increased. Additionally, those 4 variants were not involved in the ligand binding sites and the post-translational modification. So, the further research of the 4 nsSNPs should be considered as reasons of disease related to TLR3 gene malfunction.

REFERENCES

Adzhubei, I., D.M. Jordan and S.R. Sunyaev (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, 76(1): 7-20.

Akira, S., S. Uematsu and O. Takeuchi (2006). Pathogen recognition and innate immunity. *Cell*, 124(4): 783-801.

Ashkenazy, H., E. Erez, E. Martz, T. Pupko and N. Ben-Tal (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic acids Research*, 38(suppl_2): W529-W533.

Assmann, T.S., L.D.A. Brondani, A.P. Bouças, L.H. Canani and D. Crispim (2015). Toll-like receptor 3 (TLR3) and the development of type 1 diabetes mellitus. *Archives of endocrinology and metabolism*, 59(1): 4-12.

Blom, N., S. Gammeltoft and S. Brunak (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites1. *Journal of Molecular Biology*, 294(5): 1351-1362.

Botos, I., D.M. Segal and D.R. Davies (2011). The structural biology of Toll-like receptors. *Structure*, 19(4): 447-459.

Bromberg, Y., G. Yachdav and B. Rost (2008). SNAP predicts effect of mutations on protein function. *Bioinformatics*, 24(20): 2397-2398.

Bugge, M., B. Bergstrom, O.K. Eide, H. Solli, I.F. Kjønstad, J. Stenvik, T. Espevik and N.J. Nilsen (2017). Surface Toll-like receptor 3 expression in metastatic intestinal epithelial cells induces inflammatory cytokine production and promotes invasiveness. *Journal of Biological Chemistry*, 292, 15408-15425.

Capriotti, E., P. Fariselli and R. Casadio (2005). I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic acids Research*, 33(suppl_2): W306-W310.

Choi, Y. and A.P. Chan (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, 31(16): 2745-2747.

De Baets, G., J. Van Durme, J. Reumers, S. Maurer-Stroh, P. Vanhee, J. Dopazo, J. Schymkowitz and F. Rousseau (2011). SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants. *Nucleic acids research*, 40(D1): D935-D939.

Fernandez-Escamilla, A.M., F. Rousseau, J. Schymkowitz and L. Serrano (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nature Biotechnology*, 22(10): 1302

Johansen, M.B., L. Kiemer and S. Brunak (2006). Analysis and prediction of mammalian protein glycation. *Glycobiology*, 16(9): 844-853.

Johansson, M.U., V. Zoete, O. Michielin and N. Guex (2012). Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. *BMC Bioinformatics*, 13(1): 173.

Kawai, T. and S. Akira (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunology*, 11(5): 373.

Kozakov, D., L.E. Grove, D.R. Hall, T. Bohnuud, S.E. Mottarella, L. Luo, B. Xia, D. Beglov and S. Vajda (2015). The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nature protocols*, 10(5): 733.

Kumar, P., S. Henikoff and P.C. Ng (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols*, 4(7): 1073.

Maurer-Stroh, S., M. Debulpaep, N. Kuemmerer, M.L. De La Paz, I.C. Martins, J. Reumers, K.L. Morris, A. Copland, L. Serpell, L. Serrano and J.W. Schymkowitz (2010). Addendum: Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nature Methods*, 7(10): 855.

Nakamura, M., K. Funami, A. Komori, T. Yokoyama, Y. Aiba, A. Araki, Y. Takii, M. Ito, M. Matsuyama, M. Koyabu and K. Migita (2008). Increased expression of Toll-like receptor 3 in intrahepatic biliary epithelial cells at sites of ductular reaction in diseased livers. *Hepatology International*, 2(2): 222-230.

Pandey, S. and D.K. Agrawal (2006). Immunobiology of Toll-like receptors: emerging trends. *Immunology and Cell Biology*, 84(4): 333-341.

Petersen, B., T.N. Petersen, P. Andersen, M. Nielsen and C. Lundegaard (2009). A generic method for assignment of reliability scores applied to solvent accessibility predictions. *BMC Structural Biology*, 9(1): 51.

Radivojac, P., V. Vacic, C. Haynes, R.R. Cocklin, A. Mohan, J.W. Heyen, M.G. Goebl and L.M. Iakoucheva (2010). Identification, analysis, and prediction of protein ubiquitination sites. *Proteins: Structure, Function, and Bioinformatics*, 78(2): 365-380.

Salaun, B., I. Coste, M.C. Rissoan, S.J. Lebecque and T. Renno (2006). TLR3 can directly trigger apoptosis in human cancer cells. *The Journal of Immunology*, 176(8): 4894-4901.

Salaun, B., S. Lebecque, S. Matikainen, D. Rimoldi and P. Romero (2007). Toll-like receptor 3 expressed by melanoma cells as a target for therapy?. *Clinical Cancer Research*, 13(15): 4565-4574.

Schymkowitz, J., J. Borg, F. Stricher, R. Nys, F. Rousseau and L. Serrano (2005). The FoldX web server: an online force field. *Nucleic acids research*, 33(suppl_2): W382-W388.

Studzińska, M., A. Jabłońska, M. Wiśniewska-Ligier, D. Nowakowska, Z. Gaj, Z.J. Leśnikowski, T. Woźniakowska-Gęsicka, J. Wilczyński and E. Paradowska (2017). Association of TLR3 L412F Polymorphism with Cytomegalovirus Infection in Children. *PloS one*, 12(1): e0169420.

Szklarczyk, D., A. Franceschini, M. Kuhn, M. Simonovic, A. Roth, P. Minguez, T. Doerks, M. Stark, J. Muller, P. Bork and L.J. Jensen (2010). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic acids research*, 39(suppl_1): D561-D568.

Tanabe, H., Y. Fujii, M. Okada-Iwabu, M. Iwabu, Y. Nakamura, T. Hosaka, K. Motoyama, M. Ikeda, M. Wakiyama, T. Terada and N. Ohsawa (2015). Crystal structures of the human adiponectin receptors. *Nature*, 520(7547): 312.

Tassia, M.G., N.V. Whelan and K.M. Halanych (2017). Toll-like receptor pathway evolution in deuterostomes. *Proceedings of the National Academy of Sciences*, 114(27): 7055-7060.

Van Durme, J., S. Maurer-Stroh, R. Gallardo, H. Wilkinson, F. Rousseau and J. Schymkowitz (2009). Accurate prediction of DnaK-peptide binding via homology modelling and experimental data. PLoS Computational Biology, 5(8): .e1000475.

Yin, S. and B. Gao (2010). Toll-like receptor 3 in liver diseases. *Gastroenterology Research and Practice* Volume 2010, Article ID 750904, 6 pages.

Zhang, Y. and J. Skolnick (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic acids research*, 33(7): 2302-2309.

(*Accepted for publication October 2019*)