EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Volume: 8
Issue: 6
427-438
Year: 2021

IN SILICO ANALYSIS FOR SPECULATING EFFECTS OF THE DELETERIOUS SNPs OF HUMAN *CD2AP* GENE ON ITS STRUCTURE AND FUNCTIONS ASSOCIATED WITH NEPHROTIC SYNDROME

Glory S. Parmar[1], Kinnari N. Mistry[1]*, Sishir Gang[2], Dharamshibhai N. Rank[3] and Chaitanya G. Joshi[4]

[1]Ashok and Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences, Vallabh Vidyanagar, 388121, Anand, Gujarat, India.
[2]Muljibhai Patel Urological Hospital, Dr. V.V. Desai Road, Nadiad, 387001, Gujarat, India.
[3]Department of Animal Breeding and Genetics, College of Veterinary Sciences and Animal Husbandry, Anand Agricultural University, Anand 388110, Gujarat, India.
[4]Department of Animal Biotechnology, College of Veterinary Sciences and Animal Husbandry, Anand Agricultural University, Anand 388110, Gujarat, India.

***Corresponding Author: Dr. Kinnari N. Mistry**
Ashok and Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences, Vallabh Vidyanagar, 388121, Anand, Gujarat, India.

Article Received on 09/04/2021 **Article Revised on 30/04/2021** **Article Accepted on 20/05/2021**

ABSTRACT
Nephrotic syndrome is a kidney disorder that is clinically characterized by proteinuria, hyperlipidemia, hypoalbuminemia, and edema. This study aims to discover nsSNPs of the *CD2AP* gene through various *in silico* tools. A computational application helps us to reduce genotyping costs. *CD2AP* has a key role in the slit diaphragm assembly and function. Data of *CD2AP* was retrieved from NCBI dbSNP and Uniport KB and further used to explore adverse outcomes using SIFT, PolyPhen, SNP&GO, PHD SNP, PANTHER, PROVEAN SNAP2, I-Mutant, Mupro, and Strum. Out of 438 nsSNPs selected in this study, 11 SNPs: E287D, L156R, L288R, R111C, R500C, D16G, W309G, L127P, G23R, D125V, and W308S were having a damaging effect in all selected computational tools. Consurf, Hope, SOPMA, and GeneMANIA online tools were utilized to detect the evolutionary conservation of amino acids, structural impact of mutation, prediction of the secondary structure of the protein, and PPI network for functional annotation of a gene. The 3D structure was generated using I-TASSER and the stability of the built structure was confirmed through PROCHECK to get a Ramachandran plot. Post translation modification sites were checked by using Modpred software and ligand binding sites were detected using COACH. Molecular Docking analysis showed that prednisolone has the highest binding affinity towards receptor molecules than that of cyclophosphamide and levamisole. Prednisolone might be a good treatment option. The present study demonstrates the harmful effect of Non-synonymous SNPs found in the *CD2AP* gene, which alters the amino acid interactions responsible for the change in its protein function.

KEYWORDS: *CD2AP* gene; In-silico tools; Focal segmental glomerulosclerosis (FSGS); Nephrotic syndrome; Protein-Protein Interaction (PPI).

Abbreviations: *CD2AP*, CD2 associated protein; NS, Nephrotic syndrome; FSGS, Focal segmental glomerulosclerosis; PPI, Protein-Protein Interaction; nsSNP, non-synonymous single nucleotide polymorphisms; ESKD, End stage renal diseases; PTM, Post Translation Modifications; ACE, Atomic constant energy, SD, Slit diaphragm.

1. INTRODUCTION
CD2AP is a gene that encodes for CD2 associated protein, a cytoplasmic ligand of the CD2 receptor on natural killer cells. *CD2AP* is an 80KD protein expressed in all tissue that accepts the brain. It is highly expressed in kidney glomeruli.[1] It is an adaptor protein that interacts with several membranes and signaling molecules.[2] It has a vital role in regulating the actin cytoskeleton.[3] The CD2 associated protein plays a role in actin remodeling via synaptopodin binding. Knockdown of *CD2AP* gene in mice results in heavy proteinuria shortly after birth, which draws attention towards the significant role of *CD2AP* gene in glomerular function. Some glomerular lesions were observed in mice having a similar phenotype to that of human Focal Segmental Glomerulosclerosis (FSGS).[4] Kim and his colleague reported a total of 30 African American patients with idiopathic FSGS and 15 patients with HIV-associated FSGS with abnormal *CD2AP* gene function. This study is suggesting the *CD2AP* gene as a

determinant of human susceptibility to glomerular diseases.[5] Genetic detection of *CD2AP* causes nuclear relocation of dendrins, resulting in the extension of mesangial volume and mesangial fibronectin driven from a dendrin regulated podocyte secreted factor. Nuclear dendrin causes proteinuria, effacement of foot process and damage to podocyte leads to renal failure and death of the animal was observed.[6]

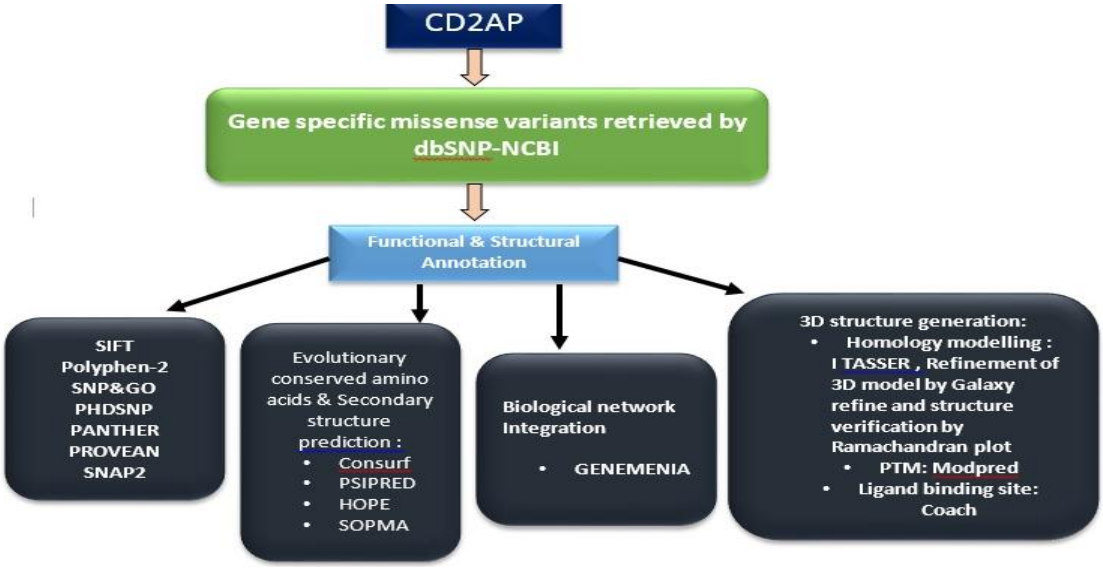


Fig. 1: Work Flow of Computational analysis of CD2 Associated protein.

One of the most common causes of heavy proteinuria in kidney disease is FSGS which progresses to End stage renal diseases (ESKD).[7] Mutations in the *CD2AP* gene have been observed in a patient having FSGS.[8] *CD2AP* gene mutation causes autosomal dominant and autosomal recessive nephrotic syndrome. It is also reported that heterozygous mutation in the *CD2AP* gene is responsible for nephrotic syndrome in adults with autosomal dominant patterns of inheritance, leading to FSGS.[9] A three-and-a-half-year-old boy with steroid-resistant nephrotic syndrome and FSGS histopathology was caused by a heterozygous mutation in the *CD2AP* gene combined with the *NPHS2* variant.[10] *In Silico* analysis of genes helps in identifying the damaging effect of amino acid substitution; which is responsible for the change in structure and function of protein. The present study aimed to check the effect of nsSNP on the structure and function of the protein, which may involve the pathogenesis of the Nephrotic syndrome (NS).

2. MATERIALS AND METHODS

Data collection: SNP data of *CD2AP* gene procured from the NCBI dbSNP database, the uniport database (uniport KB ID Q9Y5K6). The overall workflow of *In Silico* analysis is shown in **Fig.1**.

2.1 Tools for speculating the functionality of a protein

An amino acid sequence of the *CD2AP* gene was procured from the online database NCBI in FASTA format. The functional study of SNPs was carried out using seven computational tools such as SNP&GO, PolyPhen-2, SIFT, SNAP2, PROVEAN, PHDSNP and PANTHER. **SIFT:** This tool uses sequence homology and score assessment is based on specific scoring matrices. When sift score is less than 0.05, it is considered disease-causing. Whereas the score is above 0.05, it is considered tolerated.[11]

PolyPhen-2: This tool speculates the possible effects of amino acid change occurring in missense SNPs on the function and the structure of the human protein.[12]

SNP&GO: This tool helps us to predict disease-causing SNPs with a scoring efficiency of 82%.[13] Based on multiple alignments, it develops scores by calculating probabilities of all possible amino acid substitutions. When a score is >0.05, then it is considered as tolerated and score < 0.05 than it is considered as disease-causing.[14]

PANTHER: This tool also predicts the damaging effect of missense SNPs. It calculates the substitution position-specific evolutionary conservation (sub-PSEC) score to estimate whether amino acid substitution will cause any functional changes using the Hidden Markov model (HMM).

PROVEAN: This tool speculates the effect of amino acid change on the biological function of the protein. It categories the SNP variants into two types a) deleterious and b) Neutral with 79.5% accuracy rate.[15]

PHD SNP: This is an online functionality tool based on support vector machine classifiers; it classifies mutations into disease-causing and neutral polymorphism based on the protein sequence. A probability greater than 0.5 is considered disease-related whereas: a value below 0.5 is considered neutral.[16]

SNAP2: SNAP2 speculates the functional effect of amino acid change with an 83% accuracy rate, reduces runtime, and allows cross-genome comparison.[17]

2.2 Tools for predicting structural stability of protein

I Mutant: It is a neural network based tool which is used to speculate alteration and stability of protein due to amino acid substitution in a protein sequence. It calculates the DDG value of the mutated protein and native protein to speculate the stability of a protein.[18]

Mu-Pro: This tool is used to detect SNP mutations and destabilization of the protein structure. It is based on support vector machine and neural networks. It shows 84%accuracy, Mu-pro does not require a tertiary structure of a protein to predict its stability.[19] A score of less than 0 indicates the stability of protein decreases, whereas a score greater than 0 indicates protein stability increases.

STRUM: The stability of protein upon single amino acid change can be checked using strum. To construct input in Model-I of STRUM, the FASTA format of normal protein was altered at all the given positions of point mutations with substituted amino acids.[20]

2.3 Biophysical and physiochemical properties of CD2 Associated protein:

SOPMA (https://npsaprabi.ibcp.fr/cgibin/npsa_automat.pl?page=/%20NPSA/npsa_sopma.html) is an online software it uses five unique algorithms to speculate the secondary structure of the protein. (Geourjon and Deleage, 1995). PSIPRED (http://bioinf.cs.ucl.ac.uk /psipred/) is a web server utilized to predict and validates the secondary structure of *CD2AP*.[21] HOPE (https://www3.cmbi.umcn.nl/hope/input/) is an online service used to analyze the structural effects of mutation of interest.[22] The FASTA format sequence of *CD2AP* protein was retrieved from the uniport database. FASTA sequence is used as input query to check the biophysical and physiochemical properties of CD2 associated protein.

2.4 Evolutionary conservation analysis of nsSNPs by Consurf and network interaction by GeneMANIA

Consurf: This tool estimates the evolutionarily conserved amino acid/nucleic acid position in a protein/RNA/DNA molecule gleaned from the phylogenetic relations between homologous sequences.[23] A unique algorithm was used to calculate its conservation score. The conservation score between seven to nine shows that amino acids were evolutionary conserved.

GeneMANIA: It is a web server that helps to investigating gene lists and emphasizing genes for functional assay. It is an effective tool for any biologist because of its huge database and its high level of accuracy of the algorithm.[24]

Table 1: Detection of protein stability in human CD2-associated protein.

Sr No.	rsID	SNP	I-Mutant	STRUM	MuPro
1	rs147947745	E287D	Decreases	Decreases	Decreases stability
2	rs750946190	L156R	Decreases	Decreases	Decreases stability
3	rs752029188	L288R	Decreases	Decreases	Decreases stability
4	rs752979654	R111C	Decreases	Decreases	Decreases stability
5	rs758166745	R500C	Decreases	Decreases	Decreases stability
6	rs771797788	D16G	Decreases	Decreases	Decreases stability
7	rs778481750	W309G	Decreases	Decreases	Decreases stability
8	rs779581496	L127P	Decreases	Decreases	Decreases stability
9	rs1215288604	G23R	Decreases	Decreases	Decreases stability
10	rs1338314912	D125V	Decreases	Decreases	Decreases stability
11	rs1360128185	W308S	Decreases	Decreases	Decreases stability

Table 2: *CD2AP* gene functions and its association in network and genome.

Function	FDR	Genes in network	Genes in genome
Epidermal growth factor receptor signaling pathway	0.205232248	4	213
Glomerular epithelial cell differentiation	0.205232248	2	13
Glomerular epithelium development	0.205232248	2	13
Regulation of cell shape	0.205232248	3	69
ERBB signaling pathway	0.205232248	4	216
Renal filtration cell differentiation	0.205232248	2	12
Regulation of cell morphogenesis	0.205232248	4	236
Leukocyte migration	0.205232248	4	231
Positive regulation of vascular endothelial growth factor receptor signaling pathway	0.205232248	2	13
Glomerular visceral epithelial cell differentiation	0.205232248	2	12

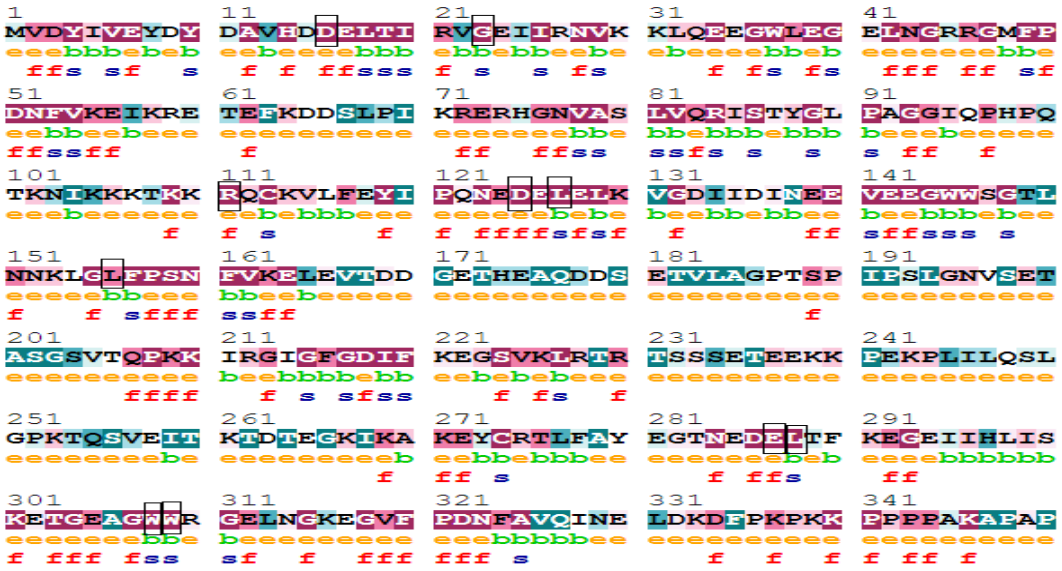
2.5 3D structure generation, visualization, and post-translation modification sites of CD2 associated protein: Homology modelling refers to the construction of the 3D structure of a novel or mutated protein (target protein) based on the known structure of a protein resembling it (template protein). For Homology Modelling, I-TASSER software was used. It is a hierarchical protein structure modelling approach based on secondary structure enhanced Profile-Profile threading Alignment (PPA) and the iterative implementation of the Threading Assembly Refinement (TASSER) program.[25] After the 3D structure generation, the Galaxy tool was used to refine the protein model. It improves the quality of structure, both global and local, on the average model generated by 3D structure prediction servers.[26] Validation of generated model: The selected model is then processed for stability assessment using PROCHECK to get a Ramachandran plot**.** Post translation modification sites were identified by ModPred. It is a sequenced-based predictor for Post Translation modification (PTM) sites in a protein sequence[27] and ligand binding sites were identified by COACH.

2.6 Molecular Docking: Molecular Docking was done by using PatchDock. This docking tool is based on the principles of shape complementarity.[28] The PatchDock algorithm separates the Connolly dot surface presentation of the protein molecules into three types: a) convex, b) concave and c) flat patches.[29]

PDB co-ordinate file of the protein and ligand molecule were used as the input parameters for the docking analysis. PatchDock analysis followed three major steps they are a) surface patch matching, b) Molecular shape representation and c) filtering and scoring.

3. RESULTS

In the present study, data of the *CD2AP* gene were collected from dbSNP (NCBI), a widely used database. It accommodates total 36874 SNPs; out of which 438 were missense SNPs, 178 were synonymous SNPs, 34601 were intronic SNPs and remaining are other types of SNPs. For our investigation, 438 missense SNPs were selected.



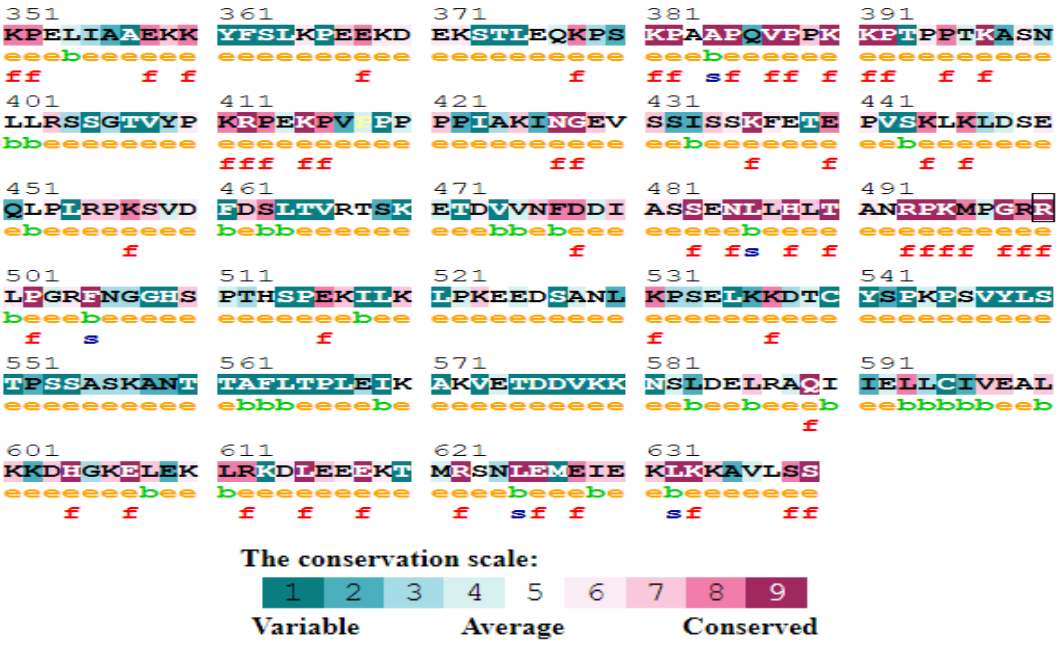


Fig 2: Evolutionary conservancy of *CD2AP* generated by consurf.

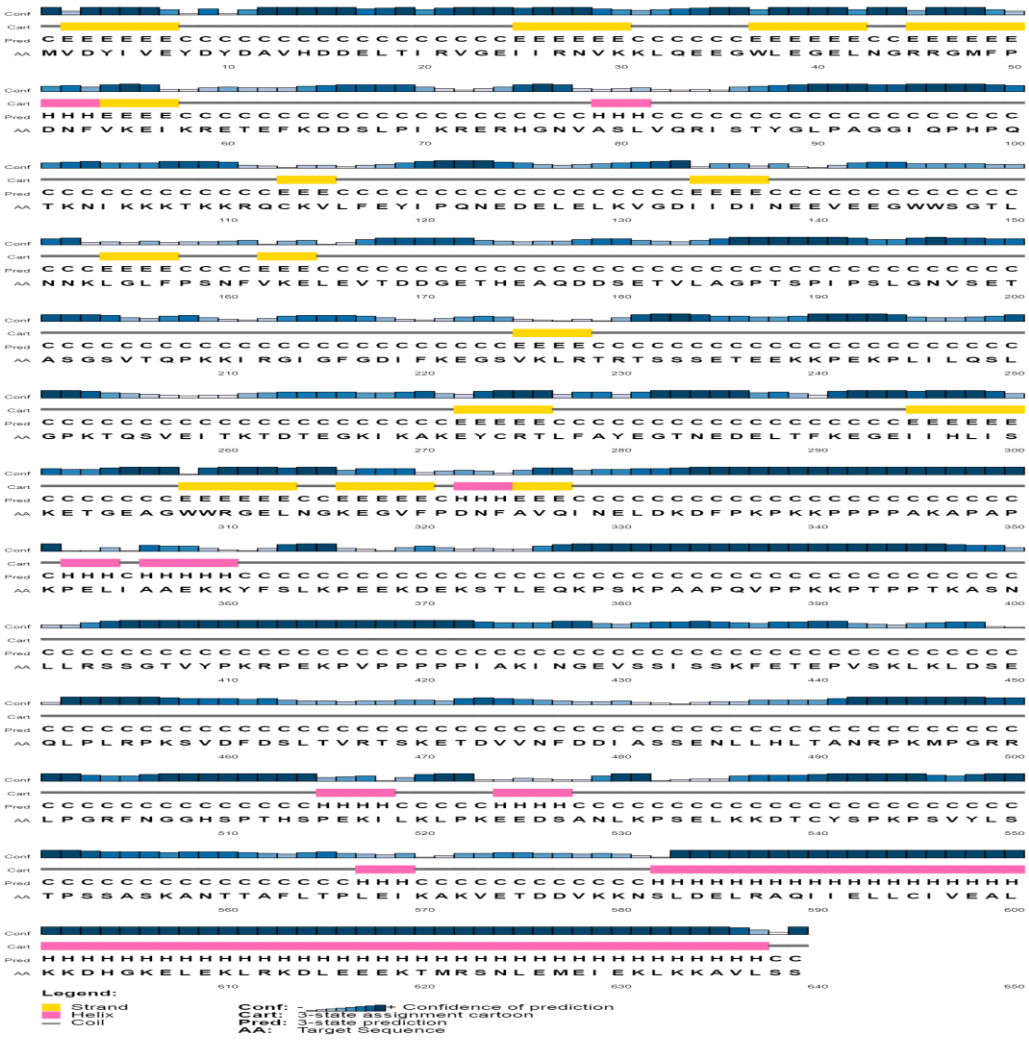


Fig 3: Secondary Structure prediction of CD2 Associated protein by PSIPRED.

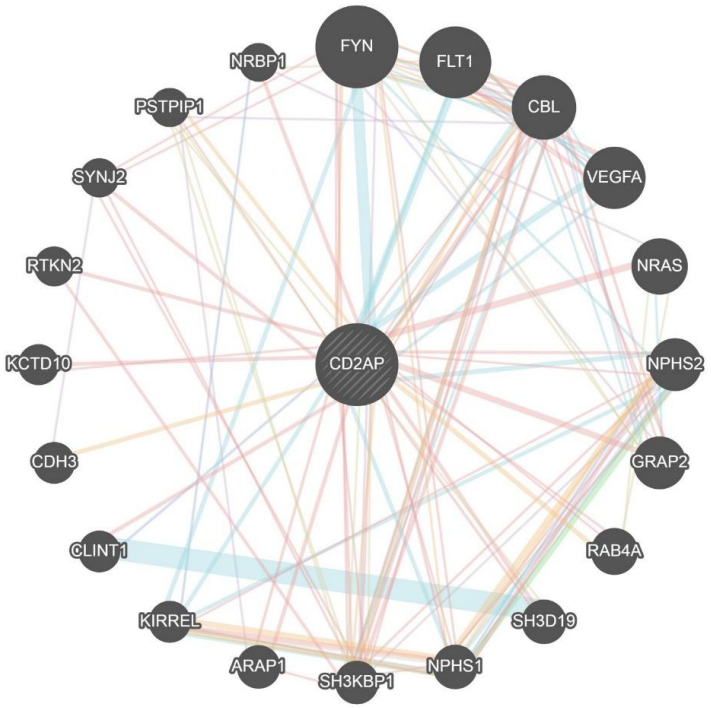


Fig.4: Biological network interaction between CD2AP and its related genes.

3.1 Functional Analysis

SIFT: The impact of any change in amino acid properties can be resolved by using SIFT by aligning orthologous protein and homologous sequence. SIFT tool was utilized for the analysis of total 438 nsSNP variants of the *CD2AP* gene. It speculates disease causing variants with good accuracy from the database. SIFT predicts nsSNPs that affect its index score (If score ≤ 0.05, it predicts as damaging and if the score is >0.05 than it predicts as tolerated). Out of 438 nsSNP variants it detected 222 SNPs as having damaging effect. These 222 missense SNPs were additionally scrutinized by using other functional tools. FASTA sequence of *CD2AP* protein and its amino acid substitutions was used as input parameter in SIFT tool. Output of SIFT analysis shows effect on protein function and SIFT score (Affect protein function or Tolerated). The tolerant index is inversely proportional to the functional impact of any amino acid change. The high tolerance index of nsSNPs has a lower functional effect and vice versa.[14] **PolyPhen-2:** PolyPhen-2 further analyzed 438 nsSNPs, it predicted 222 nsSNPs as having probably and possibly damaging effect. It predicts the amino acids causing the structural changes with 82% of accuracy. It uses BLAST algorithm to recognize the homologous protein for every input variant. The score was classified into three types 1) benign it score ranges from 0-0.9, 2) possibly damaging, score ranges from 1.4-1.9 and 3) probably damaging it scores is 2.0 or greater than this value.[30] Out of 222 missense SNPs, 137 SNPs were prognosticated as probably damaging. **SNAP2:** This tool predicts 140 SNPs as having a damaging effect on the function of a protein. It is rely on machine e- learning device known as neural network. It discriminates between neutral and effect SNP variant by considering the heterogeneity of variant and sequence characteristics. Its prediction accuracy is about 82%. **SNP&GO:** SNPs predicted by SNAP2 were further confirmed using SNP&GO. It predicts 110 SNPs having a harmful effect and the rest other SNPs having a neutral effect**. PHDSNP:** Results of SNP&GO were further confirmed using the PHDSNP tool**.** PHDSNP predicts 70 SNPs as disease-causing SNPs rest other SNPs having a neutral effect. **PROVEAN and PANTHER:** PANTHER predicts 112 SNPs as probably damaging and PROVEAN predicts 110 SNPs as probably damaging SNPs. We short-listed those nsSNPs which are common in at least seven of these above algorithm tools. Total 44 non-Synonymous SNPs showing pathogenic effect using five computational tools. So from the result of functional analysis of *CD2AP* protein, we selected 44 non-Synonymous SNPs for structural analysis of this protein: E312K, E287D, K55T, R111H, D322V, L586R, G215R, L156R, L288R, R111C, C274F, E126K, R500C, P50T, G47A, V54I, Y280C, G213E, G217D, D16G, F157L, G36E, G47R, P158A, W309G, G315C, G311C, L127P, N323Y, G44R, G23R, G318V, D478E, W308R, G40R, D473G, N77Y, D125V, W308S, G44A, G318R, Y8C, D322Y, T19I.

3.2 Protein structure Stability of CD2 associated protein: Structural Analysis

In the present study, I-Mutant, Mu-Pro and strum tools were utilized to check the stability of CD2 associated protein. Out of 44 SNPs submitted for the stability analysis we found 11 nsSNPs: E287D, L158R, L288R, R111C, R500C, D16G, W309G, L127P, G23R, D125V, and W308S showed a decrease in protein stability in all three structure stability tools and other variants showed an increase in protein stability. **(Table 1)** Consequently,

these missense variants are responsible for protein damage by disturbing its stability. It is also reported that if protein stability decreases, it causes certain harmful changes like protein degradation, misfolding, and aggregation.

3.3 Evolutionary Conservation Analysis and Secondary protein structure Analysis

Here we used the Consurf tool to analyze the evolutionary conserved CD2 associated protein Fig.3. According to the consurf tool analysis all these 11 nsSNPs: E287D, L158R, L288R, R111C, R500C, D16G, W309G, L127P, G23R, D125V, and W308S located in a highly conserved region having score between 8-9 shown in **Fig.2.** For estimation of the secondary structure of CD2 associated protein SOPMA, online software was utilized. The protein is composed of 639 amino acids having 23.32% (149 amino acids in alpha helix, 9.08% (58 amino acids) in an extended strand, 6.42% (41 amino acids) in beta-turn, and 61.91% (391 amino acid in a random coil. The graphical representation of the secondary structure of *CD2AP* by PSIPRED is shown in **Fig.3**. (Out of 11 nsSNPs, 2 SNPs: W308S and W309G were located strand part while rest of other nsSNPs were located in coil region).

3.4 Biological Network interaction by GeneMENIA

The biological functions of the *CD2AP* gene were analyzed with the GeneMANIA tool. From the annotation information recorded in GeneMANIA, a protein-protein interaction network involved with the *CD2AP* gene is shown in **Fig.4**. It showed the role of a gene in different functions like 1) Glomerular epithelial cell differentiation, 2)Glomerular epithelium development, 3)Regulation of cell shape, 4)ERBB signaling pathway, 5)Renal filtration cell differentiation, 6)Regulation of cell morphogenesis, 7)Leukocyte migration, 8)Positive regulation of vascular endothelial growth factor receptor signaling pathway, 9)Glomerular visceral epithelial cell differentiation **(Table 2).**

3.5 Structural Analysis

Effect of mutation on the structure of CD2 associated protein by Project HOPE

HOPE predicted that all 11 ns SNPs have changes in Size, charge, hydrophobicity and are located in a highly conserved position. The modification of the chemical and physical properties of amino acids upon these pathogenic mutations causes alteration in protein structure and interactions between molecules and domains were lost, leading to changes in the protein function **(Fig.5).**

3.6 Protein modelling, post-translation modification, and docking analysis.

The 3D structure of the CD2-associated protein was developed by I-TASSER. In this tool total, five models were generated; out of these five models, the best model is selected based on its c-score. *CD2AP* Structure was visualized in pymol**.** And validation of protein structure was done by the PROCHECK server **(Fig.6).** Out of 639 residues, 437 (80.2%) were in the most favored regions, 83 (15.2%) were in additionally allowed regions, 10 (1.8%) were in generously allowed regions and 15(2.8%) residues were in disallowed regions.

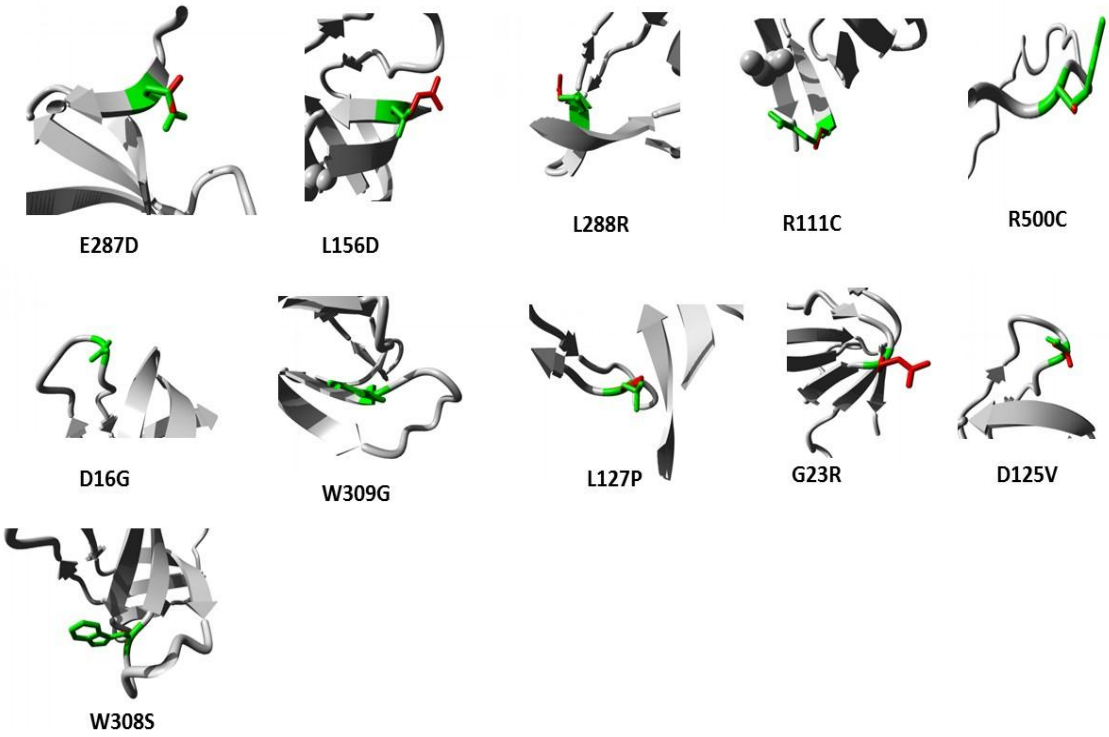


Fig.5: Structural variation of the normal residues by abnormal residues demonstrated by Project Hope. Normal residue is represented in green color and abnormal residue is represented in red colour.

Table 3: Post Translation Modification sites in *CD2AP* protein.

Sr No	rsID	SNP	Post translation modification sites Modpred
1	rs147947745	E287D	Proteolytic cleavage
2.	rs758166745	R500C	Proteolytic cleavage and Methylation

The event of PTM plays a major role in maintaining protein stability, protein localization, regulating enzyme activity and maintaining the protein's structural stability.[31] Modpred identified a site for proteolytic cleavage at E287D and R500C **(Table 3).**

Ligand binding site was predicted by the COACH server. Based on two different methods: a) S-site and b) TM-site[32], results are shown in **Table 4**. (The residues marks with bold values are included in the present study). From molecular docking analysis, we found that out of selected three drugs, the highest binding affinity was observed in prednisolone drug than that of cyclophosphamide and levamisole drugs **(Fig.7.) (Table 5).**

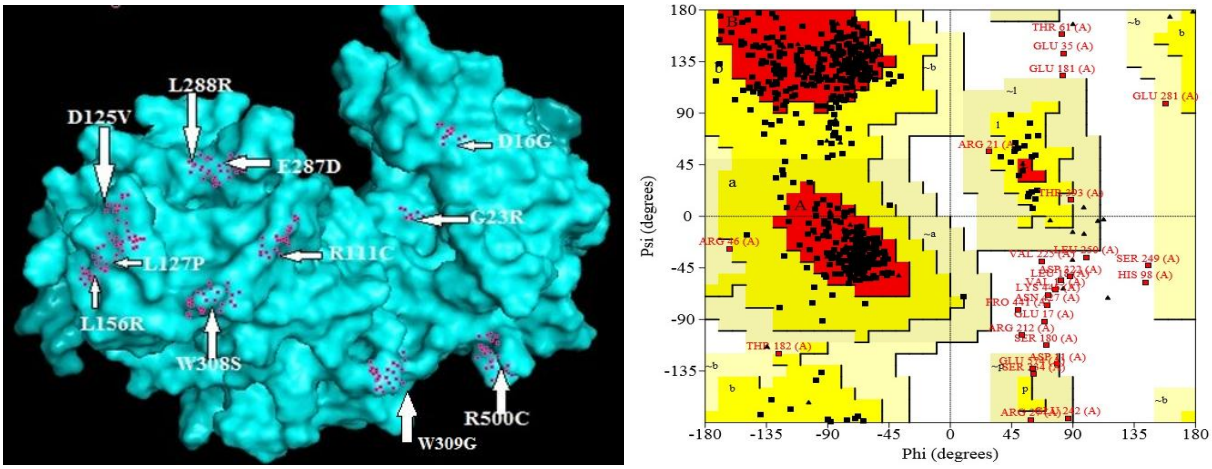


Fig.6: 3D structure generation of CD Associated protein by I TASSER and validation through Ramachandran plot analysis (Out of 639 residues, 437 (80.2%) were in the most favoured regions).

4. DISCUSSION

CD2 associated protein is a component of slit diaphragm and has crucial role in podocyte function. Limited data were available regarding the *CD2AP* gene in human renal pathology. It is actin binding protein which is essential for cytoskeletal regulation.[33] Its expression is seen at the podocyte foot process and has an important role in maintaining the integrity of the SD by interacting with other protein molecules like *NPHS1*, *NPHS2*.[4] Besides this, it also play crucial role in assembling SD complex, cell signaling pathways, *CD2AP* variants may cause deprivation in endocytosis and vesicle trafficking leading to enlarge vulnerability to toxic injury and podocyte damage.[34] *In silico* analysis can enable in detecting the phenotypic outcome of missense variants on the physio-chemical properties of the protein. This information helps out to understand the role of the *CD2AP* gene in a disease condition. Therefore attempt was made to point out deleterious SNPs in *CD2AP* gene. The single nucleotide polymorphisms are presumed to perform significant role in various human disease. Around more than 4 million human SNPs have been recorded by several SNP-related databases such as the National Center for Biotechnology Information (NCBI), Human genome variation database (HGV Database), dbSNP (database of single nucleotide polymorphisms).[35] In human about 2% of nsSNP variants are linked with genetic diseases in exonic regions, leading to functional variations.[36]

Table 4: Identification Binding site in CD-2 associated protein COACH.

COACH Results			
C-Score	**Cluster size**	**Name of Ligand**	**Binding Residues**
0.11	6	PEPTIDE	117,**125**,136,144,145,**156**,158,160,161
0.07	4	PEPTIDE	8,10,14,**16**,17,34,34,35,37,48,50,52,53
0.05	3	ZMK	582,583
0.04	2	XE	18,19,20,27,**127**
0.04	2	CL	18,126,**127**,131
0.04	2	SE	621,622
0.02	1	XE	**287**,**288**,289,294,296,297

0.02	1	CA	626,630
0.02	1	CL	130,157
0.02	1	ZN	589,592
TM-SITE RESULTS			
C- Score	Cluster size	Name of Ligand	Binding Residues
0.14	3	ZMK,PNS,SDO	582,583
0.12	2	SE,THJ	621,622
0.10	1	MG	272,533
0.09	1	ZN	589,592
0.09	1	DY	613,616
S-SITE RESULTS			
C- Score	Cluster size	Name of Ligand	Binding Residues
0.17	5	III(5)	117,119,122,123,**125**,126,141,144,145,154,**156**, 158,160,161
0.15	4	III(4)	8,10,14,**16**,17,34,35,36,37,48,50,52,53
0.10	1	AZI(1)	126,145
COFACTOR Results			
C- Score	TM -score	Name of Ligand	Binding Residues
0.01	0.354	CL	130,157
0.01	0.358	XE	**287**,**288**,289,294,296,297
-0.01	0.354	CL	18,126,**127**,131
-0.01	0.354	OAA	4,**127**
-0.01	0.358	XE	18,19,20,27,**127**
ConCavity Results			
c-Score	Binding Residues		
0.44	41,42,43,45,59,117,119,121,123,124,134,135,149,150,151,152,154,213,214,215,217,218		
0.43	8,**16**,17,18,27,39,41,126,143,624		
0.23	118,119,120,121,134,136,146,155,**156**,157,159,160,161		

Around 500000 SNPs fall in the coding region of the human genome having great significance.[37] Among this, missense SNPs are having an important role in human diseases as it accounts for a change in single amino acid residue that can lead to a functional change in a protein.[38] SNPs that are responsible for functional changes can have a damaging or neutral effect on their protein structure.[39] Alteration in protein structure was due to the pathogenic effect of variants[40], Harmful effect also responsible for altering protein hydrophobicity and charge[41] and it also disturbs inter/intra protein connections[42]; therefore cell structure integrity is lost.[43] Based on the above reasons, we can say that nsSNPs play a key role in many human diseases. Literature shows 50% of mutations are accountable for various genetic disorders[44] including many inflammatory and autoimmune disorders.[45, 46] It is well proven that CD2AP integrity is needed for the steady functioning of the slit diaphragm and the structural abnormalities involving this protein are linked with proteinuria. Experiments in mice with a targeted disruption of CD2AP[4] reveals the importance of this protein. show the significance of this protein. In vivo studies shows that loss of function of the *CD2AP* gene results in apoptotic cell death.[47]

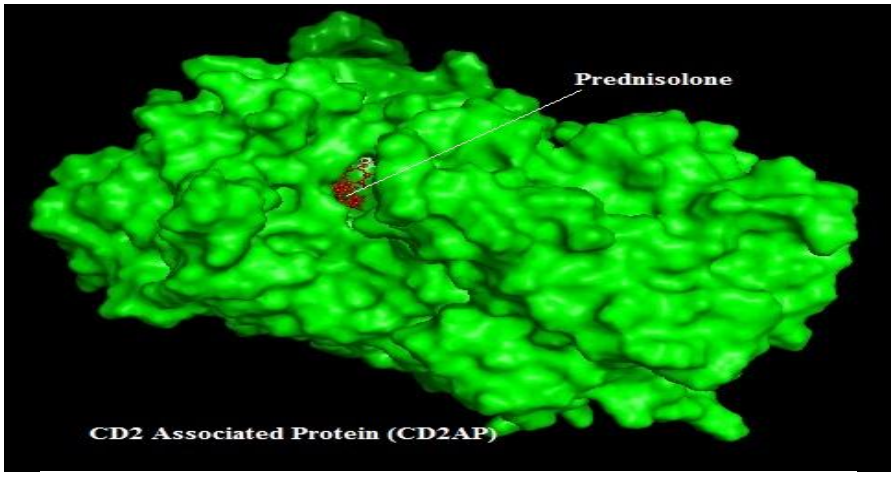


Fig.7: Molecular Docking of CD2 Associated Protein with Prednisolone.

Table 5: Molecular docking of CD2 associated proteins with steroid and non-steroid drugs ACE: Atomic Constant Energy.

CD2 Associated protein		
Drug Name	**Score**	**ACE**
prednisolone	4900	-22.59
Cyclophosphamide	4062	-138.53
Levamisole	3742	-256.98

E287D (**rs147947745)** due to this mutation there is size difference between wild type and mutant residue was observed, which might lead to loss of interaction. This mutation is located within a domain, annotated as SH3 3, which can interrupt its function. The wild type resides located in a highly conserved region. **L156R (rs750946190)** here we observed change in size as well charge difference between wild type and mutant residue. The size of mutant residue is bigger and has positive charge than wild type residue. The mutation is observed within a domain, annotated as SH3 2; due to differences in amino acid properties, it results in abnormal function of a gene. This mutant residue is located near a highly conserved position. **L288R** (**rs752029188**) the mutant residue is bigger in size and has a positive charge than that of the wild type. The mutation is observed within a domain annotated as SH3 3; due to a change in amino acid properties, the interaction was lost between other molecules and can hinder its function. This mutation is found in the highly conserved region. **R111C (rs752979654)** the mutation of arginine into a cysteine at position 111. The mutant residue is smaller and has neutral charges and the wild type has a positive charge. Due to loss of charge this mutation may cause weak interaction with other molecules. It is found within a domain; annotated as SH3 2. This mutation disturb the function of the protein due to loss of interaction between other important molecules. It is found in highly conserved region. **R500C (rs758166745)** this mutation reduces the size of normal residue. The charge of mutant residue is neutral and the wild type has a positive charge**.** The charge of the wild type is lost due to this mutation which hinders the interaction with other molecules. This mutation is found on the surface of the domain having some unknown function. **D16G (rs771797788)** the size of this mutant residue is smaller compared to its wild-type residue. Mutant residue has a neutral charge and wild type has a negative charge. Change in the size as well as charge responsible for loose interaction with other molecules. This mutation is situated within a domain; annotated as SH3 1 due to a change in properties of amino acid; that can interrupt this domain interaction with other molecules and abolish its function. This mutation is found in the highly conserved region.

W309G (rs778481750) the mutation of tryptophan into a glycine at position 309. The mutant residue is smaller than wild-type residue, this might lead to loss of interactions and hydrophobic interaction either in the core of the protein or the surface will be lost. The mutation is located within a domain, annotated as SH3 3 due to a change in the properties of amino acids; can interrupt this domain and terminate its functions. Glycine is very flexible; due to this mutation the rigidity of the protein at this position is lost. This mutation is located in a highly conserved region. **L127P (rs779581496)** the mutation of leucine into a proline at position 127. The mutant residue is smaller than wild type residue. The mutation will cause space in the core of the protein. It is located within a domain annotated as SH3 2, due to a change in the properties of amino acids which can lead to a change in its function and disturb the interaction between domain and other molecules. HOPE predicted that other residues that have the same properties in common with this mutation were observed. This suggests that in some rare cases L127P mutation might occur without damaging protein. This mutation occurs in a highly conserved region. **G23R (rs1215288604)** the mutant residue is bigger than the wild type and has a positive charge while wild type residue has a Neutral charge. Project hope predicted that due to change in charge of the wild type residue there was problem in

protein folding process. This mutation is located within a domain annotated as SH3 1, due to this mutation flexibility of the protein is lost which can lead to abnormal function of the protein. **D125V (rs1338314912)** in this mutation the charge of wild type is lost. It is found within a domain; annotated as SH3 2, due to a change in properties of amino acids the interaction between this domain is lost and it abolishes its function. This mutation occurs in a highly conserved region. **W308S (rs1360128185)** the mutation of tryptophan into a serine at position 308. The mutant residue is smaller which might lead to loss of interactions. The hydrophobic interaction either in the core of the protein or on the surface will be lost. This mutation is located within a domain, annotated as SH3 3. Project Hope predict this mutation as probably damaging based on its conservation information. The present study demonstrated the role of deleterious SNPs of CD2AP gene by using In Silco based approach.

5. CONCLUSION

Computational based studies are very beneficial in genetic association study as it minimizes the cost of SNP genotyping. Presents study identified a total of 11 Non-Synonymous/missense SNPs in *CD2AP* gene: E287D, L156R, L288R, R111C, R500C, D16G, W309G, L127P, G23R, D125V, and W308S having a pathogenic effect in all ten functional and structural computational tools. So, from our result, we can conclude that these Non-Synonymous SNPs might be accountable for the pathogenesis of NS. This study will help narrow down SNP variants for experimental studies that might be involved in disease-causing. The result of molecular docking showed that the drug prednisolone having the highest binding affinity towards selected podocyte genes. Because of its antioxidant activity it might be one of the good treatment options to manage NS. Therefore, an *in silico* based study is the most convenient approach to know the impact of DNA variation on the function of a gene.

Conflicts of Interest

No conflict exits.

ACKNOWLEDGMENT

Authors are immensely thankful to CharutarVidya Mandal Vallabh Vidyanagar. We would also like to thank GSBTM (Gujarat State Biotechnology Mission) Gandhinagar, for providing financial assistance to carry out this research work (GSBTM/FAP 2016 -Project ID No. 1390).

REFERENCES

1. Kim JM, Wu H, Green G, et al. CD2-associated protein haploinsufficiency is linked to glomerular disease susceptibility. Science, 2003; 300(5623): 1298-300.
2. Lehtonen S, Ora A, Olkkonen VM, et al. In vivo interaction of the adapter protein CD2-associated protein with the type 2 polycystic kidney disease protein, polycystin-2. J Biol Chem, 2000; 275(42): 32888-93.
3. Kirsch KH, Georgescu M-M, Ishimaru S, Hanafusa H. CMS: an adapter molecule involved in cytoskeletal rearrangements. Proc Natl Acad Sci, 1999; 96(11): 6211-6216.
4. Shih N-Y, Li J, Karpitskii V, et al. Congenital nephrotic syndrome in mice lacking CD2-associated protein. Science, 1999; 286(5438): 312-315.
5. Gigante M, Pontrelli P, Montemurno E, et al. CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). Nephrol Dial Transplant, 2009; 24(6): 1858-1864
6. Gigante M, Pontrelli P, Montemurno E, et al. CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). Nephrol Dial Transplant, 2009 ; 24(6): 1858-1864.
7. Weins A, Wong JS, Basgen JM, et al. Dendrin ablation prolongs life span by delaying kidney failure. Am J Pathol, 2015; 185(8): 2143-2157.
8. Gbadegesin RA, Winn MP, Smoyer WE. Genetic testing in nephrotic syndrome—challenges and opportunities. Nat Rev Nephrol, 2013; 9(3): 179-184.
9. Kaplan JM, Kim SH, North KN, et al. Mutations in ACTN4, encoding α-actinin-4, cause familial focal segmental glomerulosclerosis. Nature Genet, 2000; 24(3): 251-256.
10. Löwik M, Levtchenko E, Westra D, et al. Bigenic heterozygosity and the development of steroid-resistant focal segmental glomerulosclerosis. Nephrol Dial Transplant, 2008; 23(10): 3146-3151.
11. Kumar A, Purohit R. Computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E. MUTAT RES-FUND MOL M, 2012; 738: 28-37.
12. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat. Methods, 2010; 7(4): 248-249.
13. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. HUM MUTAT, 2009; 30(8): 1237-1244.
14. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res, 2003; 31(13): 3812-3814.
15. Choi Y, Sims GE, Murphy S, Miller JR, and Chan A.P,. Predicting the functional effect of amino acid substitutions and indels. PloS one, 2012; 7(10): 1-13.
16. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. BMC Genom, 2013; 14(3): 1-7.
17. Yachdav G, Hecht M, Pasmanik-Chor M, Yeheskel A, Rost, B. HeatMapViewer: interactive display of 2D data in biology. F1000Res, 2014; 3(48): 1-6.

18. Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res, 2005; 33(2): W306-W310.
19. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins, 2006; 62(4): 1125-1132.
20. Quan L, Lv Q, Zhang Y. STRUM: structure-based prediction of protein stability changes upon single-point mutation. Bioinformatics, 2016; 32(19): 2936-2946
21. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics, 2000; 16(4): 404-405.
22. Wang Z, Huang C, Lv H, Zhang M, Li X, In Silico analysis and high-risk pathogenic phenotype predictions of non-synonymous single nucleotide polymorphisms in human Crystallin beta A4 gene associated with congenital cataract. Plos one, 2020; 15(1): 1-19.
23. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 2016; 44(W1): W344-W350.
24. Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res, 2010; 38(2): W214-W220.
25. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinform, 2008; 9(1): 1-8.
26. Wang L, Jin Y, Arnoldussen YJ, et al. STAMP1 is both a proliferative and an antiapoptotic factor in prostate cancer. Cancer Res, 2010; 70(14): 5818-5828.
27. Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, Radivojac P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci, 2014; 23(8): 1077-1093.
28. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res, 2005; 33(2): W363-W367.
29. Doss CGP, Sudandiradoss C, Rajasekaran R, Choudhury P, Sinha P, Hota P, Batra, UP, Rao S. Applications of computational algorithm tools to identify functional SNPs. Funct Integr Genomics, 2008; 8(4): 309-316.
30. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res, 2002; 30(17): 3894-3900.
31. Maeda A, Okano K, Park PS-H, et al. Palmitoylation stabilizes unliganded rod opsin. Proc Natl Acad Sci, 2010; 107(18): 8428-8433.
32. Yang J, Roy A, Zhang Y. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics, 2013; 29(20): 2588-2595.
33. Lehtonen S, Zhao F, Lehtonen E. 2002. CD2-associated protein directly interacts with the actin cytoskeleton. Am J Physiol Renal Physiol, 2002; 283(4): F734-F743.
34. Wolf G, Stahl RA. CD2-associated protein and glomerular disease. The Lancet, 2003; 362(9397): 1746-8.
35. Rajasekaran R, Sudandiradoss C, Doss CGP, Sethumadhavan R, Identification and In Silico analysis of functional SNPs of the BRCA1 gene. Genomics, 2007; 90(4) 447-452.
36. Fredman D, Siegfried M, Yuan YP, Bork P, Lehväslaiho H, Brookes AJ. HGVbase: a human sequence variation database emphasizing data quality and a broad spectrum of data sources. Nucleic Acids Res, 2002; 30(1): 387-391.
37. Collins FS, Brooks LD, Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. Genome Res, 1998; 8(12): 1229-1231.
38. Lander ES. The new genomics: global views of biology. Science, 1996; 274(5287): 536-539.
39. Capriotti E, Altman RB. Improving the prediction of disease-related variants using protein three-dimensional structure. BMC Bioinform, 2011; 12(4): 1-11.
40. Barroso I, Gurnell M, Crowley V, et al. Dominant negative mutations in human PPARγ associated with severe insulin resistance, diabetes mellitus and hypertension. Nature, 1999; 402(6764): 880-883.
41. Petukh M, Kucukkal TG, Alexov E. On human disease-causing amino acid variants: Statistical study of sequence and structural patterns. Hum Mutat, 2015; 36(5): 524-34.
42. Kucukkal TG, Petukh M, Li L, Alexov E. Structural and physico-chemical effects of disease and non-disease nsSNPs on proteins. Curr Opin Struct Biol, 2015; 32: 18-24.
43. Thomas R, McConnell R, Whittacker J, Kirkpatrick P, Bradley J, Sandford R. Identification of mutations in the repeated part of the autosomal dominant polycystic kidney disease type 1 gene, PKD1, by long-range PCR. Am J Hum Genet, 1999; 65(1): 39-49.
44. Radivojac P, Vacic V, Haynes C, et al. Identification, analysis, and prediction of protein ubiquitination sites. Proteins, 2010; 78(2): 365-380.
45. Azad AK, Sadee W, Schlesinger LS, Innate immune gene polymorphisms in tuberculosis. Infect Immun, 2012 80(10): 3343-3359.
46. Heim MH, Innate immunity and HCV. J Hepatol, 2013; 58(3): 564-574.
47. Huber TB, Hartleben B, Kim J, et al. Nephrin and CD2AP associate with phosphoinositide 3-OH kinase and stimulate AKT-dependent signaling. Mol Cell Biol, 2003; 23(14): 4917-4928.