ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

In silico analysis of functional nsSNPs of the Human *PRPS1* gene.

B Preethi and K Ramanathan*

Industrial Biotechnology Division, School of Bio Sciences and Technology, VIT University, Vellore- 632014, Tamil Nadu, India.

ABSTRACT

In the present investigation, different genomics algorithm employed for the identification of nsSNPs in *PRPS1* (phosphoribosyl pyrophosphate synthetase 1) gene. The missense mutations in *PRPS1* gene will leads to Branchio-oto-renal Syndrome (BOR Syndrome). The initiation of the investigation was done with SIFT followed by 5 other genomic algorithms such as POLYPHEN-2, I-Mutant 2.0, Mupro, mCSM and Align GVGD. A total of 16 variants of *PRPS1* gene of *Homosapiens* were retrieved from dbSNP for our analysis. The analysis showed that 6 variants (L152P, M115T, D183H, D52H, I290T and V112A) were found to be less stable and damaging. Also outputs predicted for 6 variants shows that all have clinical impact in causing the disease. We hope these results will be undoubtedly helpful for the pharmacist working in BOR Syndrome for the development of potential drug candidate.

Keywords: BOR syndrome, Missense mutation, PRPS1, nsSNPs.

**Corresponding author*

ISSN: 0975-8585

INTRODUCTION

Single nucleotide polymorphism (SNP), defined as single base changes in a DNA sequence, are account for the more abundant form of genetic variation in the human population. Although numerous SNPs are phenotypically impartial, non-synonymous SNPs (nsSNPs) frequently have deleterious impacts on protein structure and its function. Most of the SNPs in the human genome are present in the non-coding DNA consisting of 5' and 3' and translated regions (UTR) [1]. The dbSNP is used for the same and it is a public domain archive [2]. Phosphoribosyl Pyrophosphate Synthetase 1 (*PRPS1*) is an isoform of the *PRPS* gene family and is universally expressed in human tissues, including cochlea. The enzyme mediates the biochemical step significant for purine metabolism and nucleotide biosynthesis [3]. PRPS1 codes for PRS-I, which catalyzes the synthesis of phosphoribosyl pyrophosphate (PRPP) from ATP and ribose-5-phosphate [4]. The mutations in PRS-I affect vital cell functions, such as nucleic acid synthesis, energy metabolism, and cellular signaling. X-linked non-syndromic hearing loss is caused by missense mutations in *PRPS1* gene [5]. The study found that missense mutation in the *PRPS1* gene lead to a condition called hereditary Branchio-oto-renal Syndrome (BOR Syndrome). This condition is portrayed by hearing loss in male patients with *PRPS1* mutations is bilateral, moderate to significant, and can be prelingual or postlingual, progressive or non-progressive. Female carriers can have unilateral or bilateral hearing impairment. *PRPS1* gene is located from base pair 106,871,654 to basepair 106,894,256 on chromosome X. The extent that vicinity situation is concerned the discovery of deleterious SNPs is essential task for pharmacogenomics and pharmacogenetics.

We involved this work basically to perform a computational investigation of *PRPS1* gene consisting of nsSNPs and identification of possible deleterious mutation. Out of the 16 SNPs, the most deleterious SNPs which are significant in causing disease are L152P, M115T, D183H, D52H, I290T and V112A. These mutations can be a hopeful of most concern in the BOR Syndrome.

MATERIALS AND METHODS

Dataset

The dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) is used to obtain the SNPs and their related protein sequences for *PRPS1* gene (*Homosapiens*). The database provides a unique ID for every SNP; reference ID (rsIDs) for every SNP. The rsIDs give information about the particular SNP, amino acid changes, their respective positions and corresponding accessions IDs. Clicking on accessions ID delivers information regarding the protein encoded by the genes. The protein PDB ID is 2H06. There are numerous inclusive and accessible software packages and web-based services to identify the structures [6].

Analysis of functional effect of point mutations using SIFT

To identify the damaging single amino acid polymorphism SIFT (Sorting Intolerant from Tolerant) tool was employed [7]. The technique is mainly based on the evolutionary amino acid conservation in protein families. Positions are intolerant to substitution if they are more conserved and vice versa. Therefore, results are deleterious or damaging when substitutions are at well conserved positions. Multiple sequence alignment information is used for the prediction of given sequence [8]. It predicts whether the substitution is damaging or deleterious for each position of the query sequence. SIFT process consists of the following steps (a) enter the protein rsID, (b) enter amino acid substitutions, (c) note down the tolerance index >0.5 is the cut off value of tolerance index for SIFT. This is inversely proportional to the impact of amino acid substitutions i.e. higher the tolerance index lessens the impact of amino acid substitution and vice versa.

Structure analysis using PolyPhen-2

PolyPhen-2 (Polymorphism Phenotyping v2) tool is a structure and sequence prediction tool which gives the impact of amino acid substitution on structure and function of human protein [9]. A protein sequence is given as input along with mutational positions and two variants of amino acids. The score (PSIC) of the prediction is

noted. The amino acid substitutions which are probably damaging are noted and used in further tools for prediction of impact of amino acid substitution.

Analysis of protein stability using I-Mutant 2.0

It is a support vector machine based tool. This tool predicts the change in protein stability due to single point mutations [8]. Protein sequence is given as input along with amino acid substitution and position. The output is a free energy change value ($\Delta\Delta G$). Positive $\Delta\Delta G$ value infers that the protein being mutated is of higher stability and vice versa.

Stability prediction based on protein sequence using MUpro

This tool predicts protein stability changes for single point mutations on the protein sequence. The protein sequence along with amino acid substitution and position is given as input file for analysis [10]. The result shows two methods both of which should decrease stability of protein. The substitutions which decrease stability by both methods are noted and used in further prediction tools.

mCSM study

mCSM predicts the effect of mutations in protein using graph-based signatures. These encode distance patterns between atoms and are used to represent the protein residue environment and to train predictive models [11]. In this tool the protein PDB file along with the amino acid substitution and position is given as input. The output is the free energy value (ΔG). This indicates whether the mutation is stabilizing or destabilizing. The destabilizing point mutations are noted and used for further predictions.

Align- GVGD analysis

It is a freely available, web-based tool that combines biophysical characteristics of amino acids and protein multiple sequence alignments to predict where missense substitutions in genes of interest fall in a spectrum from enriched deleterious to enriched neutral [12]. Protein sequence or PDB file is given as input along with amino acid substitution and position. The scores are classified into classes. Mutations belonging to Class C65 are most likely to interfere with function of gene and are deleterious. The mutations belonging to Class C65 are distinguished in the present study.

RESULTS AND DISCUSSIONS

There are 16 missense mutations which were retrieved from dbSNP specifically Q133P, L152P, E43D, M115T, N114S, D183H, D52H, L129I, A190V, H193Q, D65N, A87T, I290T, G306R, V112A and A233P. The missense mutations were given as input into various freely available web based programs (SIFT, POLYPHEN2, I-Mutant 2.0, MUpro, mCSM and Align GVGD).

The first software for analysis used was SIFT. The mutations were submitted in SIFT program for analyzing tolerance index [7]. We found in our analysis, out of the 16 variants, 6 variants werefound to be deleterious with a tolerance index score of >.05. The results were computed and tabulated (Table 1). It was observed that out of 16 variants 6 are highly deleterious with tolerance index 0.

After SIFT, the POLYPHEN2 program [9] was employed with protein sequence having mutational position submitted as inputs. PSIC score >0.950 were found to be probably damaging, PSIC score of > 0.5 were found to be possibly damaging and the rest were found to be benign are shown in table 1.

ISSN: 0975-8585

Table 1: Deleterious nsSNP predicted by SIFT, PolyPhen-2 and I-Mutant

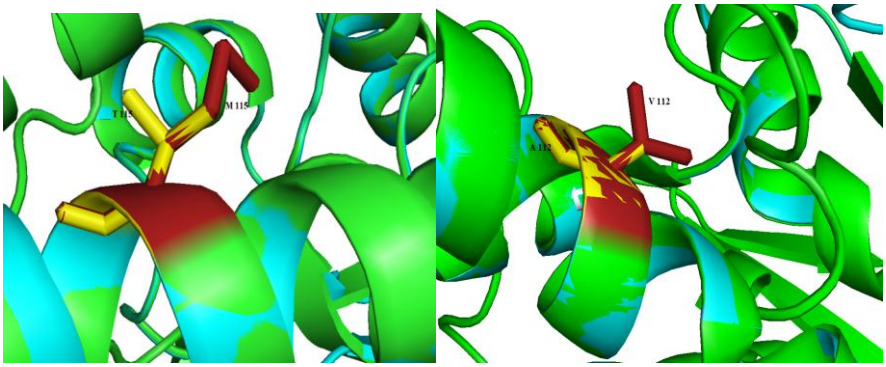
rs ID	Amino Acid change	SIFT (Tolerance index)	PolyPhen-2 (PSIC SD)	PolyPhen-2 Prediction	I- Mutant Stability
rs80338676	L152P	0	0.998	Probably Damaging	Decrease
rs80338732	M115T	0	0.950	Possibly Damaging	Decrease
rs137852541	D183H	0	1.000	Probably Damaging	Decrease
rs137852542	D52H	0	1.000	Probably Damaging	Decrease
rs180177153	I290T	0	0.538	Possibly Damaging	Decrease
rs11541075	V112A	0	0.988	Probably Damaging	Decrease

Table 2: Deleterious nsSNPs predicted by MUpro, mCSM and Align GVGD

rs ID	Amino Acid change	MUpro	mCSM (kcal/mol)	Stability Prediction	Align GVGD
rs80338676	L152P	Decrease Stability	-1.532	Destabilizing	Most likely
rs80338732	M115T	Decrease Stability	-1.133	Destabilizing	Most likely
rs137852541	D183H	Decrease Stability	-1.005	Destabilizing	Most likely
rs137852542	D52H	Decrease Stability	-1.277	Destabilizing	Most likely
rs180177153	I290T	Decrease Stability	-1.631	Destabilizing	Most likely
rs11541075	V112A	Decrease Stability	-1.755	Destabilizing	Most likely

The mutations which have a cut off >0.5 are further given as input into the program I-Mutant 2.0. This program gives ΔΔ*G* values and tells whether the amino acid substitutions increase or decrease stability (Table 1). Only the mutations which decrease the stability of the protein are taken into consideration for further studies.

Other supplementary tools have also been used to carry out the study. Tools such as MUpro, mCSM and Align GVGD are used to further confirm the deleterious nature of these mutations. The result of MUpro, mCSM and Align GVGD are shown in table 2. We found that all the three programs shows 6 variants are deleterious and destabilized. The two tables sum up the results of the study and prove the deleterious effects of mutations on PRPS1 gene. From the two tables we can infer that the given mutations are highly deleterious and will cause Branchio-oto-renal Syndrome and further these functionally significant variants were superimposed with native PRS-I protein structure (PDB ID: 2H06) using PyMol (Figure 1) [13].



(a) **(b)**

ISSN: 0975-8585

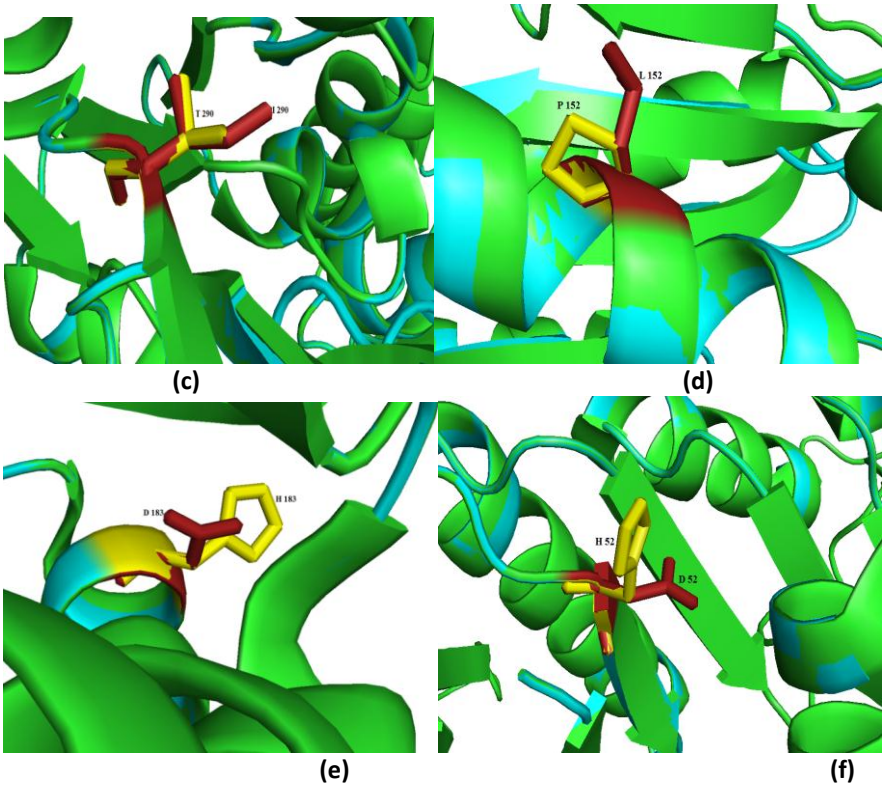


Figure 1: Superimposed images of PRS-I protein with mutant structures: (a) M115T, (b) V112A, (c) I290T, (d) L152P, (e) D183H and (f) D52H.

CONCLUSION

In the present study, we investigated the functional SNPs in *PRPS1* gene using computational methods. This study mainly aims to screen the deleterious nsSNPs in the *PRPS1* gene. It is noteworthy to mention that 6 mutations such as L152P, M115T, D183H, D52H, I290T and V112A that could impart maximum deleterious effect on *PRPS1* function. Most importantly, the majority of these high-risk nsSNPs are located at highly conserved amino acid positions. Hence, we conclude that these nsSNPs associated with *PRPS1* gene could be the important candidate for the cause of BOR Syndrome. It is certain that the above results will have immense importance in the treatment of the BOR Syndrome in the near future.

ACKNOWLEDGMENT

The authors would like to thank management of VIT University for providing the facilities to carry out this work.

REFERENCES

[1] Rajasekaran R, Sudandiradoss C, Doss CG, Sethumadhavan R. Genomics 2007; 90: 447-52.
[2] Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. Nucleic Acids Res 2001; 29: 308-11.
[3] Kim HJ, Sohn KM, Shy ME, Krajewski KM, Hwang M, Park JH, Jang SY, Won HH, Choi BO, Hong SH, Kim BJ, Suh YL, Ki CS, Lee SY, Kim SH, Kim JW. Am J Hum Genet 2007; 81(3): 552-8.
[4] Kornberg A, Lieberman I, Simms ES. J Biol Chem 1955; 215: 389–402.

[5] Liu X, Han D, Li J, Han B, Ouyang X, Cheng J, Li X, Jin Z, Wang Y, Bitner-Glindzicz M, Kong X, Xu H, Kantardzhieva A, Eavey RD, Seidman CE, Seidman JG, Du LL, Chen ZY, Dai P, Teng M, Yan D, Yuan H. Am J Hum Genet 2010; 86(1): 65-71.
[6] Kumar P, Henikoff S, Ng PC. Nat Protoc 2009; 4: 1073-81.
[7] Ng PC, Henikoff S. Nucleic Acids Res 2003; 31: 3812–14.
[8] Capriotti E, Fariselli P, Casadio R. Nucleic Acids Res 2005; 33: 306-10.
[9] Ramensky V, Bork P, Sunyaev S. Nucleic Acids Res 2002; 30: 3894–3900.
[10] Cheng J, Randall A, Baldi P. Proteins 2006; 62(4): 1125-32.
[11] Pires DE, Ascher DB, Blundell TL. Bioinformatics 2014; 30(3): 335-42.
[12] Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV. Nucleic Acids Res 2006**;** 34(5): 1317-25.
[13] Warren LD. San Carlos, CA: De Lano Scientific LLC, USA 2002.