Send Orders for Reprints to reprints@benthamscience.net

526 *Current Genomics*, 2021, *22*, 526-540

RESEARCH ARTICLE

In-Silico and *In-Vitro* Analysis of Human *SOS1* Protein Causing Noonan Syndrome - A Novel Approach to Explore the Molecular Pathways

Vinoth Sigamani[1], Sheeja Rajasingh[1], Narasimman Gurusamy[1], Arunima Panda[2] and Johnson Rajasingh[1,3,4,*]

[1]Department of Bioscience Research, University of Tennessee Health Science Center, Memphis, Tennessee; [2]Department of Genetic Engineering, SRM Institute of Science and Technology, Chennai, India; [3]Department of Medicine-Cardiology, University of Tennessee Health Science Center, Memphis, Tennessee; [4]Department of Microbiology, Immunology & Biochemistry, University of Tennessee Health Science Center, Memphis, Tennessee

Abstract: ***Aims*:** Perform *in-silico* analysis of human *SOS1* mutations to elucidate their pathogenic role in Noonan syndrome (NS).

***Background*:** NS is an autosomal dominant genetic disorder caused by single nucleotide mutation in PTPN11, *SOS1*, RAF1, and KRAS genes. NS is thought to affect approximately 1 in 1000. NS patients suffer different pathogenic effects depending on the mutations they carry. Analysis of the mutations would be a promising predictor in identifying the pathogenic effect of NS.

***Methods*:** We performed computational analysis of the *SOS1* gene to identify the pathogenic nonsynonymous single nucleotide polymorphisms (nsSNPs) th a t cause NS. *SOS1* variants were retrieved from the SNP database (dbSNP) and analyzed by *in-silico* tools I-Mutant, iPTREESTAB, and MutPred to elucidate their structural and functional characteristics.

***Results*:** We found that 11 nsSNPs of *SOS1* that were linked to NS. 3D modeling of the wild-type and the 11 nsSNPs of *SOS1* showed that *SOS1* interacts with cardiac proteins GATA4, TNNT2, and ACTN2. We also found that GRB2 and HRAS act as intermediate molecules between *SOS1* and cardiac proteins. Our *in-silico* analysis findings were further validated using induced cardiomyocytes (iCMCs) derived from NS patients carrying *SOS1* gene variant c.1654A>G (NSiCMCs) and compared to control human skin fibroblast-derived iCMCs (C-iCMCs). Our *in vitro* data confirmed that the *SOS1*, GRB2 and HRAS gene expressions as well as the activated ERK protein, were significantly decreased in NS-iCMCs when compared to C-iCMCs.

***Conclusion*:** This is the first *in-silico* and *in vitro* study demonstrating that 11 nsSNPs of *SOS1* play deleterious pathogenic roles in causing NS.

ARTICLE HISTORY

Received: August 13, 2021
Revised: November 02, 2021
Accepted: November 03, 2021

DOI:
10.2174/1389202922666211130144221

Keywords: Noonan syndrome, *SOS1* gene, *in-silico* analysis, post-translational modification, nonsynonymous SNP, pathogenic variants.

1. INTRODUCTION

Noonan syndrome (NS) is an autosomal dominant genetic disorder characterized by short stature, congenital heart disease, bleeding problems, developmental delays, and skeletal malformation. The occurrence of NS is estimated to be between 1:1000 and 1:2500 live births, and it affects more males than females [1, 2]. The molecular defects of NS are related to the altered function of RAS-MAPK signaling caused by a mutation in four main genes - *PTPN11, SOS1, RAF1*, and *KRAS*. Among these genes, the *PTPN11* is the dominant gene for NS [3, 4].

The Son of Sevenless Homolog 1 (*SOS1*) is the second dominant gene for NS, and the mutation of this gene causes a distinctive phenotype with keratosis pilaris and curly hair [5]. The substitution of Thr266Lys in *SOS1* showed facial dysmorphisms and mild pulmonic stenosis [6]. Studies have shown that in NS patients with germline mutation of *SOS1* developed tumors [7]. Individuals with NS display cardiac anomalies, such as non-syndromic pulmonic stenosis, atrial septal defects, and ventricular septal defects [8]. *SOS1* is a guanine nucleotide exchange factor, which upregulates the *RAS* signaling pathway, leading to the changes in human development [9, 10], and also alters *RAS* and *RAC1* pathways [11]. The CIIA act as a negative regulator of *RAS*-specific guanine nucleotide exchange factor activity of *SOS1* [12]. The interaction domains of *SOS1*/GRB2 control the embryonic stem cell fate during mammalian development [13].

*Address correspondence to this author at the Department of Bioscience Research, University of Tennessee Health Science Center, Memphis, Tennessee; Tel: 901-448-3358; E-mail: rjohn186@uthsc.edu

The human genome contains 1.42 million single nucleotide polymorphisms (SNPs). Among these 250,000 - 400,000 SNPs in the protein-coding region, not sequence of the genome are named as non-synonymous SNPs (nsSNPs), which alters the amino acid in their functional protein [14, 15]. About 26-32% of nsSNPs are functionally effective and lead to cause disease by changing post-translational modification (PTM), protein stability and protein-protein interaction [16, 17]. In this study, we computationally analyzed the *SOS1* gene to identify the pathogenic nsSNPs responsible for NS. The 3D models of wildtype and mutant *SOS1* proteins were analysed. Since NS is known to cause cardiac anomalies, the interactions of *SOS1* with other cardiac proteins were studied using STRING, and were experimentally validated *in-vitro* by mRNA and protein expressions in NS patient-derived induced cardiomyocytes (NS-iCMCs). This is the first *in-silico* study of the *SOS1* variants linked to NS, and to discover the molecular pathways associated with this disease.

2. MATERIALS AND METHODS

2.1. Datamining

The *SOS1* variants were retrieved from the NCBI SNP database (dbSNP) (https://www.ncbi.nlm.nih.gov/snp/) that is a public domain archive for a broad collection of single genetic polymorphisms. The protein sequence of *SOS1* was retrieved from UniProt (https://www.uniprot.org/), which provides free accessible resources of protein sequence and functional information.

2.2. Consequences of Variants

The retrieved variants were analysed by the variant effect predictor (VEP) tool (http://www.ensembl.org/Tools/VEP) to determine the likely consequences of amino acid substitutions on protein function [18].

2.3. Identifying the Most Pathogenic nsSNPs with Noonan Syndrome

The pathogenic nsSNPs of *SOS1* were filtered following a previous literature review on NS (https://www. ncbi.nlm.nih.gov/pubmed/) on NS. Further, the nsSNPs were analysed by *in-silico* tools such as dbNSFP, sorting intolerant from tolerant (SIFT), polymorphism phenotyping (PolyPhen), protein variation effect analyser (PROVEAN), functional analysis through hidden markov model (FATHMM), mutation taster, consensus deleteriousness (condel), likelihood ratio test (LRT) (https://www.ensembl.org/Homo_sapiens/Tools/VEP?db=core), single nucleotide polymorphisms and gene ontology (SNP & GO) (http://snps.biofold.org/snps-and-go/snps-and-go.html), and predictor of human deleterious single nucleotide polymorphisms (PhD-SNP) (http://snps.biofold.org/phd-snp/ phd-snp.html). Studies have shown that the SNPs were considered as more pathogenic, when it had been predicted by more than eight *in-silico* tools [18, 19].

2.4. Analysing Protein Stability

The stability of the wild type and mutant *SOS1* protein stability was predicted by I-Mutant 2.0 (http://folding.biofold.org/i-mutant/i-mutant2.0.html) and iPTREE-STAB, an interpretable decision tree-based method (http://203.64.84.190:8080/IPTREEr/iptree.htm). I-Mutant 2.0 is a support vector machine-based web server that helps in the automatic prediction of protein stability changes upon single-site mutations by using a data set derived from ProTherm. It predicts the changes in free energy delta-delta-G (DDG), which predicts how a single-point mutation affects protein stability and the results are expressed as a positive or negative value in Kcal/mol. We submitted protein sequence by changing single-site mutations of *SOS1* protein to predict the stability of proteins, while the conditions were set at the temperature 25º C and pH 7 [20, 21].

2.5. Analysis of Structural and Functional Properties of nsSNPs

The structural and functional properties of pathogenic nsSNPs were analyzed by submitting their amino acid substitution of the protein sequence (FASTA format) in MutPred2 (http://mutpred2.mutdb.org/about.html). It predicts the molecular pathogenicity of amino acid substitution and the altered molecular mechanism affecting the phenotype using the threshold P-value of ≤ 0.05 [22].

2.6. 3D Protein Modeling and Visualization

To generate the 3D models of wild type and mutants of *SOS1* protein, we submitted amino acid substitution of a protein sequence in the Iterative Threading ASSEmbly Refinement (I-TASSER) (https://zhanglab.ccmb.med.umich.edu/I-TASSER/), which is a hierarchical approach to predict the structure and function of a protein. It identifies the structural templates from the PDB, based on LOMETS approaches, and the function of the target is derived by BioLiP [23-25]. Later, these models were verified and selected based on C-Score, ERRAT score (https://servicesn.mbi.ucla.edu/ ERRAT/) [26], and also verified by Ramachandran Plot Analysis (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) [27]. Then, the verified structure was visualized on PyMOL 2.2.2, a 3D structure visualization software. The 3D structural effect on the mutation was analyzed by HOPE (http://www.cmbi.ru.nl/hope/input/), which is an automatic mutation analysis server [28]. Furthermore, to extend our structural analysis, TM-score and root mean square deviation (RMSD) between wild type and mutant was calculated using TM-align software (https://zhanglab.ccmb.med.umich.edu/TM-align/) [29].

2.7. *SOS1* Protein Network

To predict the specific interaction of *SOS1* with cardiac proteins, we submitted the protein in the STRING server (https://string-db.org/), which is a database that provides computational direct protein-protein interactions and its indirect functional associations [30].

2.8. *In-vitro* Validation of Protein Network

Finally, the findings from the *in-silico* gene interactions associated with NS were further validated by *in-vitro* conditions using induced pluripotent stem cells (iPSC)-derived cardiomyocytes (iCMCs) obtained from normal control individuals (C-iCMCs), NS patient (NS-iCMCs), and their respective parent cells; normal human skin fibroblast (C-SF), NS patient cardiac fibroblasts (NS-CF), and iPSCs derived from C-SF (C-iPSCs); NS-iCMCs, and iPSC-derived from NS-CF (NS-iPSCs) carrying *SOS1* gene variant c.1654A>G (SNP_ID.rs137852814) as demonstrated in our previous studies [31, 32].

To perform qRT-PCR, we isolated total RNA from C-SF, C-iPSCs, C-iCMCs, NS-CF, NS-iPSCs and NS- iCMCs using trizol reagent (Ambion by Life Technologies) as described in the manufacturer protocol. Then, the first strand cDNA was synthesized from each RNA sample by using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). The gene expressions for the protein-network interacted genes (*ACTN2, GATA4, GRB2, HRAS, NKX2.5, TNNT2* and *SOS1*) were performed using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems). Briefly, a total reaction volume of 10 µl was prepared with the composition of 0.5 µl of cDNA, 1 µl of 5 picomole forward and reverse primer mix, 5 µl of SYBR Green and 3.5 µl of DEPC water. The PCR amplification was carried out in a Quantstudio 6pro (Applied Biosystems) with fast thermal cycling condition, 2 minutes of UDG activation temperature at 50˚C, 2 minutes of dual-lock DNA polymerase temperature at 95˚C and followed with 40 cycles at 95˚C for 1 second and 60˚C for 30 seconds. The relative expression for the target genes was normalized with 18S rRNA as endogenous control. The results are shown in fold change expression, and the values were calculated as the ratio of induced expression to control expression. The mRNA expression was further supported by Western immunoblotting analysis of protein expression as described earlier [33]. GAPDH was used as a protein loading control. The sequence of primers used for qRT-PCR and the antibodies used for the Western immunoblotting are given in Supplementary Table **1**.

2.9. Statistical Analysis

The statistical analyses were performed between the mean of each group by Bonferroni's method of one-way ANOVA using GraphPad Prism 8 software. The P-value <0.05 is considered as statistically significant. All *in-vitro* experiments were performed with biological triplicates.

3. RESULTS

3.1. Variants of *SOS1* Genes

We used NCBI - dbSNP and UniProt were used to retrieve variants of the *SOS1* gene and the protein sequence of human *SOS1* (ID: Q07889). There was a total of 38137 variants, including 634 deletions, 3767 deletions/insertions, 11 multiple nucleotide variants and 33725 single nucleotide variants in *SOS1* gene (Fig. **1**).

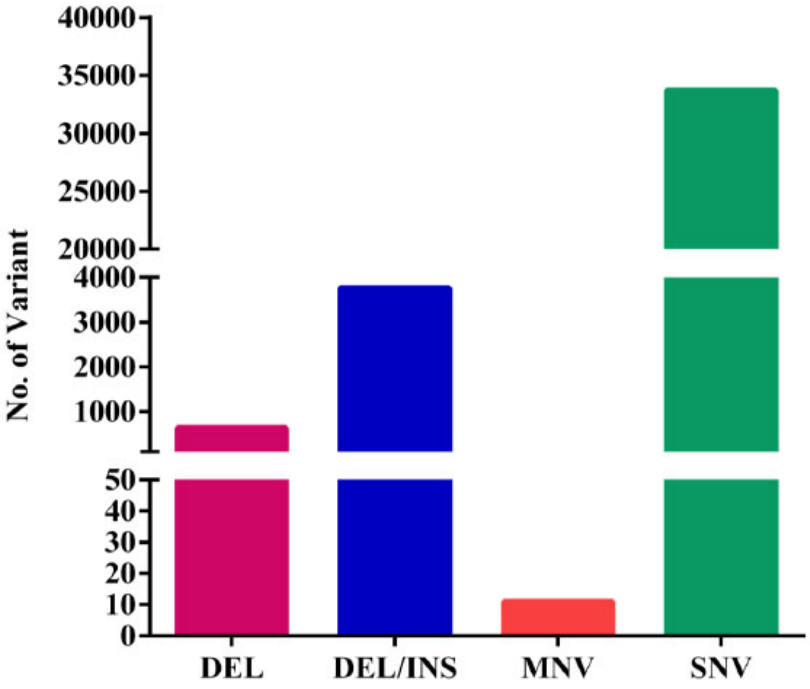

Fig. (1). Variation classes of *SOS1*_variants. Deletion (DEL) - 634, Deletion/Insertion (DEL/INS) - 3767, Multiple Nucleotide Variant (MNV) - 11, and Single nucleotide variant (SNV) - 33725. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

3.2. Consequences of *SOS1* Variants in Percentages

We have used the VEP tool to find the consequences of *SOS1* variants [18]. From a total of 38137 variants, 34058 (89.1%) occurred in the intron; 1814 (4.8%) in upstream and downstream, 1066 (2.8%) in 3' UTR, 688 (1.8%) in nonsynonymous, 320 (0.9%) in synonymous, and 191 (0.5%) in other variant classes were identified Fig. (**2**). The 688 nsSNPs were analysed further to examine their pathogenicity.

3.3. Prediction of Most Pathogenic nsSNPs in Noonan Syndrome

Based on the literature, we identified that 16 nsSNPs of *SOS1* variants were pathogenic with NS (Table **1**). The functionality of these 16 nsSNPs was predicted using *in-silico* tools. The functional prediction and scores obtained from various *in-silico* tools are given in Supplementary Table **2**. The SIFT score predicts a function of protein affected by amino acid substitution. The score ranges from 0.0 (deleterious) to 1.0 (tolerate), and the score 0.0 to 0.05 considered as deleterious. The PolyPhen score ranges from 0.0 (benign) to 1.0 (deleterious), and the results predicted to be benign, probably or possibly damaging. In the PROVEAN tool, if the score is below the threshold value of -2.5, it is considered as deleterious, and if the score is greater than the threshold, it is considered as neutral. FATHMM tool returns a *p*-value, and the variant scores ≤ 0.5 are considered deleterious. In Mutation Taster, a larger score is considered as deleterious, and based on the score ranges from 0 to 1; a variant is classified as A (disease-causing automatic), D (disease-causing), N (polymorphism), or P (polymorphism automatic). Condel tool evaluates missense single nucleotide variants and yields a score ranging from 0.0 to 1.0; the value < 0.5 is considered as neutral, and the value >0.5 is considered as deleterious. The score of LRT ranges from 0 to 1, and this can predict the SNP as damaging or neutral based on the rules described by Liu *et al*. [34]. For the SNPs & GO and PhD-SNP, the score ranges from 0.0 to 1.0; if the value of a variant is >0.5, it is predicted as disease causing, otherwise, it is considered neutral [18, 19]. Based on the results of *in-silico* analyses (Table **1** and Supplementary Table **2**), we have chosen 11 SNPs as the most pathogenic predicted based on the number of *in-silico* tools (8 out of 9 tools) and shown as deleterious (Supplementary Table **1** and Fig. **3A**) and the remaining 5 SNPs were predicted as less pathogenic (Fig. **3B**). The specificity of our analysis was confirmed with the cancer-associated *SOS1* mutation D309Y [35], which showed pathogenicity in 7 out of 9 *in-silico* tools in our study (Supplementary Table **2**).

3.4. Protein Stability Prediction

We used I-Mutant 2.0 and iPTREE-STAB tools to predict the stability of a protein resulting from the nsSNPs. Both of these tools gave a DGG value of a protein at the pH 7.0 and at the temperature at 25° C. The value of DGG can be positive or negative, and it is interpreted as the increase or decrease in protein stability, respectively [20, 21]. Our results showed that all of the 11 nsSNPs had a decreased protein stability, at least in one of the two tools (Table **2**). The protein stability scores predicted for the five less pathogenic *SOS1* mutants from I-Mutant and iPTREE are given in Table **3**.

Table 1. List of Pathogenic nsSNPs associated with NS.

SNP ID	cDNA Changes	CDS Changes	Amino Acid Changes	Disorders	References
rs137852812	c.884C>A	c.797C>A	p.Thr266Lys	NS Type 4, NS, Rasopathy, Gingival fibromatosis 1	[3-13, 45-56]
rs137852813	c.893T>G/C	c.806T>G/C	p.Met269Arg/Thr	NS Type 4, NS, Rasopathy, Inborn genetic diseases	
rs137852814	c.1741A>G	c.1654A>G	p.Arg552Gly	NS Type 4, NS, Rasopathy, Gingival fibromatosis 1	
rs267607079	c.1743G>T/C	c.1656G>C/T	p.Arg552Ser		
rs267607080	c.1381T>C	c.1294T>C	p.Trp432Arg	NS Type 4, NS	
rs397517147	c.1384G>A	c.1297G>A	p.Glu433Lys	NS, Rasopathy	
rs397517148	c.1387G>A	c.1300G>A	p.Gly434Arg	NS, Abnormality of the sternum, Ptosis, Pulmonic stenosis, Short stature, Rasopathy	
rs397517149	c.1729A>C	c.1642A>C	p.Ser548Arg	NS Type 4, NS, Rasopathy, Inborn genetic diseases	
rs397517150	c.1397T>C	c.1310T>C	p.Ile437Thr	NS, Rasopathy	
rs397517153	c.1736T>C	c.1649T>C	p.Leu550Pro		
rs397517154	c.1742G>C/A	c.1655G>C/A	p.Arg552Thr/Lys	NS, NS Type 3, Rasopathy, Abnormality of the aortic valve	
rs397517156	c.2270A>T	c.2183A>T	p.Lys728Ile	NS	
rs397517159	c.2623G>A	c.2536G>A	p.Glu846Lys	NS Type 4, NS, Rasopathy, Gingival fibromatosis 1	
rs397517164	c.409G>A	c.322G>A	p.Glu108Lys	NS, Rasopathy	
rs397517180	c.1012G>T	c.925G>T	p.Asp309Tyr	NS	
rs727504295	c.1409G>A	c.1322G>A	p.Cys441Tyr	NS, Rasopathy	

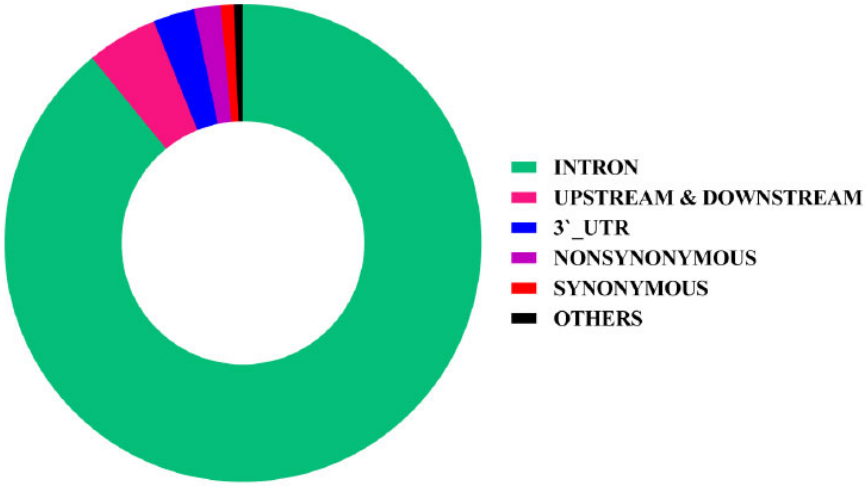

Fig. (2). Pie diagram showing the consequences of *SOS1* variants in percentages. Intron variants - 90.9%, Upstream & Downstream variants - 4.8%, 3`_UTR variants - 2.8%, nonsynonymous variants - 1.8%, Synonymous variants - 0.9%, and other variants - 0.5%. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

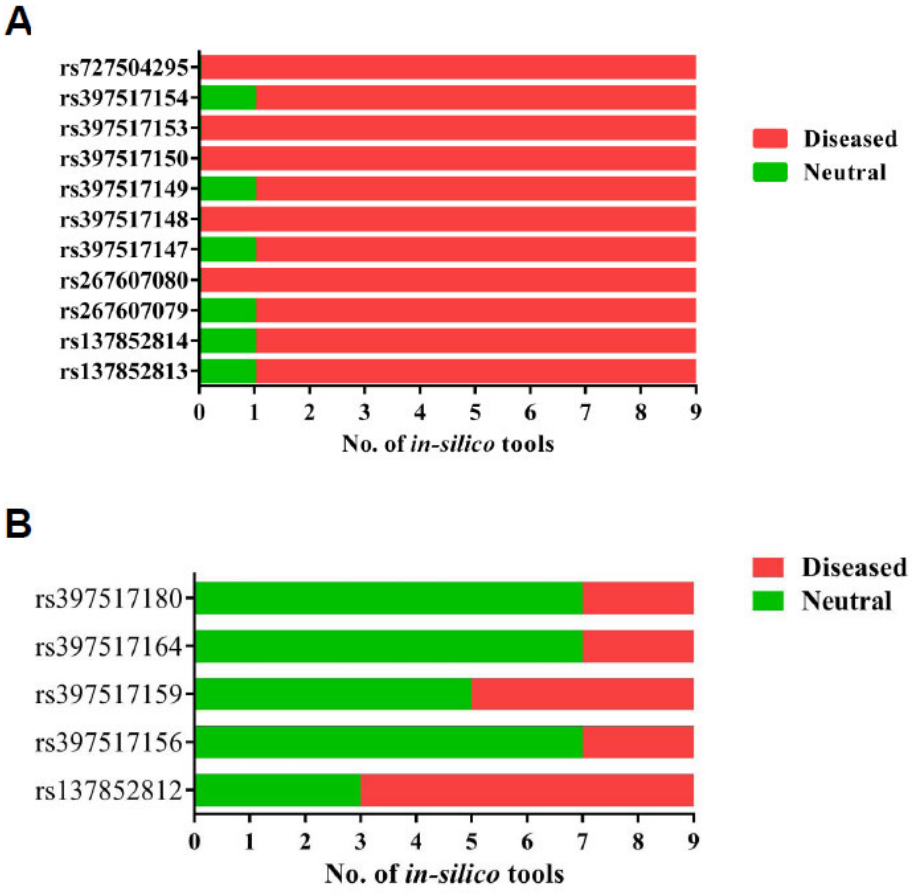

Fig. (3). The pathogenicity of *SOS1* nsSNPs predicted by various *in-silico* tools. (**A**) The most pathogenic 11 nsSNPs and (**B**) the less pathogenic 5 SNPs were associated with NS, predicted by various *in-silico* tools. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

Table 2. *SOS1* mutant proteins stability prediction by I-mutant 2.0 and iPTREE-STAB.

SNP ID	Amino Acid Changes	I-Mutant 2.0		iPTREE-STAB	
		DDG	Protein Stability	DDG	Protein Stability
rs137852813	p.Met269Arg	-0.78	Decrease	-2.46	Decrease
rs137852814	p.Arg552Gly	-1.17	Decrease	-0.60	Decrease
rs267607079	p.Arg552Ser	-2.14	Decrease	-5.10	Decrease
rs267607080	p.Trp432Arg	-1.82	Decrease	-0.11	Decrease
rs397517147	p.Glu433Lys	-0.81	Decrease	-1.85	Decrease
rs397517148	p.Gly434Arg	-1.79	Decrease	-1.77	Decrease
rs397517149	p.Ser548Arg	-0.26	Decrease	-1.55	Decrease
rs397517150	p.Ile437Thr	-2.15	Decrease	-5.10	Decrease
rs397517153	p.Leu550Pro	-0.08	Decrease	-0.60	Decrease
rs397517154	p.Arg552Thr	-1.12	Decrease	-5.10	Decrease
rs727504295	p.Cys441Tyr	0.83	Increase	-0.02	Decrease
rs137852812	p.Thr266Lys	-0.28	Decrease	0.1725	Increase
rs397517156	p.Lys728Ile	0.27	Increase	-1.8289	Decrease
rs397517159	p.Glu846Lys	-0.82	Decrease	-1.3925	Decrease
rs397517164	p.Glu108Lys	-0.41	Decrease	-0.6080	Decrease
rs397517180	p.Asp309Tyr	-0.77	Decrease	-0.4600	Decrease

Table 3. The protein stability scores predicted for the five less pathogenic *SOS1* mutants by I-mutant 2.0 and iPTREE-STAB.

SNP ID	Amino Acid Changes	I-Mutant		iPTREE-STAB	
		DDG	Protein Stability	DDG	Protein Stability
rs137852812	p.Thr266Lys	-0.28	Decrease	0.1725	Increase
rs397517156	p.Lys728Ile	0.27	Increase	-1.8289	Decrease
rs397517159	p.Glu846Lys	-0.82	Decrease	-1.3925	Decrease
rs397517164	p.Glu108Lys	-0.41	Decrease	-0.6080	Decrease
rs397517180	p.Asp309Tyr	-0.77	Decrease	-0.4600	Decrease

Table 4. Prediction of altered molecular mechanisms of *SOS1* nsSNPs by mutPred2.

SNP ID	Amino Acid Changes	MutPred2 Score	Molecular Mechanisms with *P*-Values <= 0.05		
			Prediction	Probability	*P*-Value
rs137852813	p.Met269Arg	0.782	Gain of Intrinsic disorder	0.3	0.05
rs137852814	p.Arg552Gly	0.464	-	-	-
rs267607079	p.Arg552Ser	0.447	-	-	-
rs267607080	p.Trp432Arg	0.896	Gain of Intrinsic disorder	0.41	0.0075
			Gain of Acetylation at K427	0.3	0.0039
			Altered Coiled coil	0.28	0.02
rs397517147	p.Glu433Lys	0.528	Altered Coiled coil	0.53	0.004
			Gain of Helix	0.27	0.04
rs397517148	p.Gly434Arg	0.573	Gain of Helix	0.28	0.02
			Altered Coiled coil	0.25	0.02
rs397517149	p.Ser548Arg	0.349	-	-	-
rs397517150	p.Ile437Thr	0.644	Altered Metal binding	0.23	0.04
rs397517153	p.Leu550Pro	0.849	Loss of Helix	0.33	0.0016
			Loss of Proteolytic cleavage at D555	0.21	0.0017
			Altered Transmembrane protein	0.16	0.01

(Table 4) contd....

SNP ID	Amino Acid Changes	MutPred2 Score	Molecular Mechanisms with *P*-Values <= 0.05		
			Prediction	Probability	*P*-Value
rs397517154	p.Arg552Thr	0.519	Loss of Helix	0.29	0.01
			Loss of Proteolytic cleavage at D555	0.2	0.0024
			Altered Transmembrane protein	0.15	0.01
rs727504295	p.Cys441Tyr	0.922	Loss of Helix	0.33	0.0016
			Gain of Relative solvent accessibility	0.28	0.02
			Gain of Strand	0.27	0.02
			Altered Metal binding	0.26	0.02

Table 5. Structural verification scores for the wild type and selected mutant models of *SOS1*.

SNP ID	Amino Acid Changes	C - Score	ERRAT Value	Ramachandran Plot			TM-Score	RMSD (Å)
				Favoured Region	Allowed Region	Outlier Region		
Wild Type	-	-2.4	82.459	78.40%	12.50%	9.10%	-	-
rs137852813	p.Met269Arg	-2.49	96.641	75.40%	12.70%	11.90%	0.7823	0.55
rs137852814	p.Arg552Gly	-2.23	85.832	75.40%	12.50%	12.10%	0.7765	0.47
rs267607079	p.Arg552Ser	-2.5	85.146	77.20%	11.30%	11.50%	0.9451	1.29
rs267607080	p.Trp432Arg	-2.16	84.783	77.50%	12.90%	9.60%	0.7856	0.49
rs397517147	p.Glu433Lys	-2.15	84.633	78.10%	11.40%	10.40%	0.9448	1.37
rs397517148	p.Gly434Arg	-2.11	81.559	75.80%	12.80%	11.30%	0.7776	0.48
rs397517149	p.Ser548Arg	-2.21	84.621	77.80%	10.40%	11.80%	0.7821	0.48
rs397517150	p.Ile437Thr	-2.34	85.008	78.30%	11.80%	9.90%	0.9465	1.26
rs397517153	p.Leu550Pro	-2.35	86.702	79.10%	10.80%	10.10%	0.9319	1.25
rs397517154	p.Arg552Thr	-1.95	81.095	75.40%	11.60%	13.10%	0.7709	0.51
rs727504295	p.Cys441Tyr	-2.67	84.441	76.40%	13.20%	10.40%	0.7864	0.5

3.5. Structural and Functional Properties of nsSNPs

The structural and functional properties of nsSNPs were evaluated by MutPred2, which predicted the probability of deleterious mutations and the alterations in molecular mechanism, if it obtains a P-value of ≤ 0.05 [22]. Our results showed that eight nsSNPs, listed in Table **4**, have deleterious mutations with altered mechanisms with a P-value of ≤ 0.05. However, three nsSNP (IDs rs137852814, rs267607079, and rs397517149) had a score of <0.5 and classified as 'tolerated' by MutPred2.

3.6. 3D Modelling and Visualization of *SOS1* Protein

We have generated 3D protein models for the wild type and 11 nonsynonymous mutants of *SOS1* by using I-TASSER, which produced five models for each variant [23-25]. These models were verified and one model was for each variant based on the model having a minimum C-score, maximum ERRAT score, and the most allowed region on Ramachandran plot [26, 27]. The results of the selected models are shown in Table **5**. Furthermore, the chosen models of wild type and mutants were visualized by PyMOL 2.2.2 software; and the HOPE server [28] was used to mask their amino acid substitution (Fig. (**4** and Supplementary Fig. **1**). To extend our structural analysis, we calculated template modeling score (TM-score) and root-mean-square deviation of atomic positions (RMSD) for the 11 nonsynonymous mutant models, and compared them with wild type models [29]. The TM-score shows the topological similarity between wild type and mutant models. All the mutants had a score of > 0.5, which indicated that the models were similar. The RMSD value measures the average distance between the α-carbon backbones of the wild type and mutant models. The higher RMSD value indicated a greater deviation between the wild type and mutant models. The results of TM-align, along with its RMSD score, are shown in Table **5**.

3.7. Interactions of *SOS1* with Cardiac-Specific Proteins

The interactions of *SOS1* with cardiac proteins were predicted by the STRING server [30], which predicted the physical and functional interactions of proteins. *SOS1* was found to interact with many cardiac proteins, such as ACTN2, ACTN4, GATA4, NKX2.5, TNNI3 and TNNT2, mainly through GRB2 and HRAS (Fig. **5**).

3.8. *In-vitro* Validation Using NS-iCMCs in Comparison with C-iCMCs

We verified the mRNA and protein expressions of *SOS1* interacting cardiac proteins in NS-iCMCs and compared those levels with C-iCMCs. qRT-PCR analyses have revealed that the mRNA expression of *ACTN2, GATA4, TNNT2* (Fig. **6A**) and *GRB2, HRAS, SOS1* (Fig. **6B**) were significantly decreased, and *NKX2.5* (Fig. **6C**) was significantly increased in NS-iCMCs compared with C-iCMCs. To support these mRNA expression profiles, we performed Western blot analyses, which showed that GATA4, GRB2, HRAS, and *SOS1* proteins and ERK1/2 activation were significantly reduced in NS-iCMCs compared with C-iCMCs (Fig. **6D**).

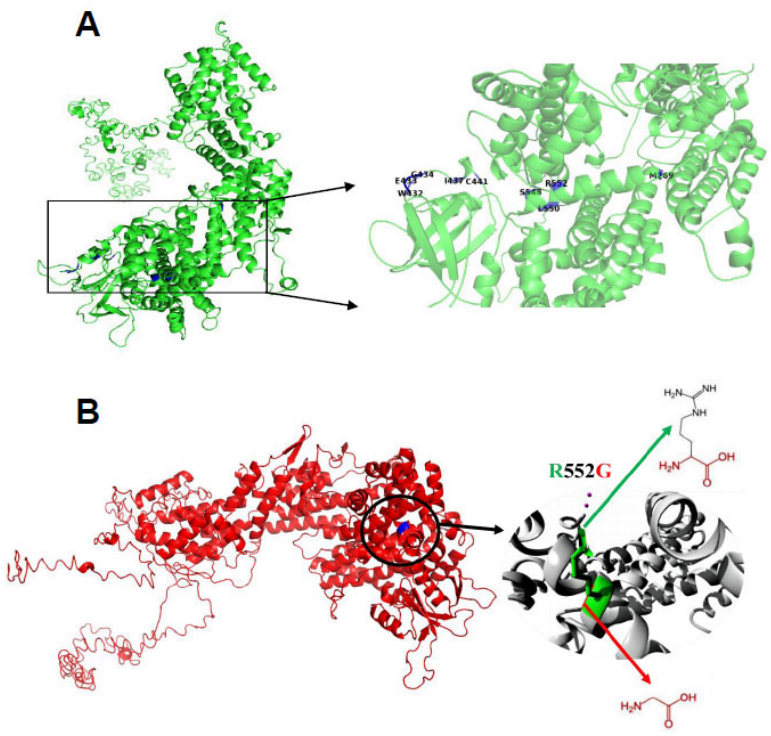

Fig. (4). 3D model of *SOS1* wild type and its R552G mutant. **(A)** Green, wild type model of *SOS1* and **(B)** Red, R552G mutant model of *SOS1* (rs137852814) showing an amino acid change at the 552nd position from Arginine to Glycine. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

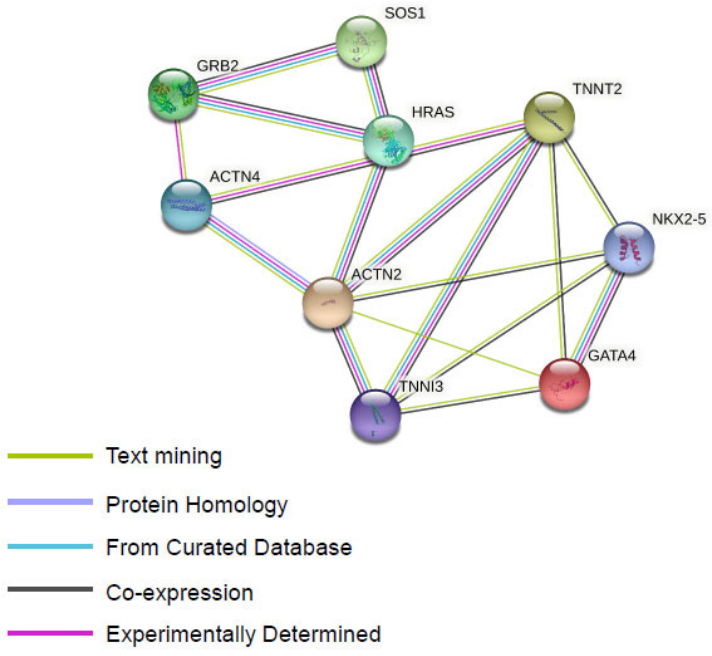

Fig. (5). Interactions of *SOS1* with cardiac-specific proteins. The interactions of *SOS1* with cardiac proteins were predicted using STRING server, which showed that *SOS1* interacts with GATA4, TNNT2, TNNI3, ACTN2, ACTN4 through GRB2 and HRAS. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

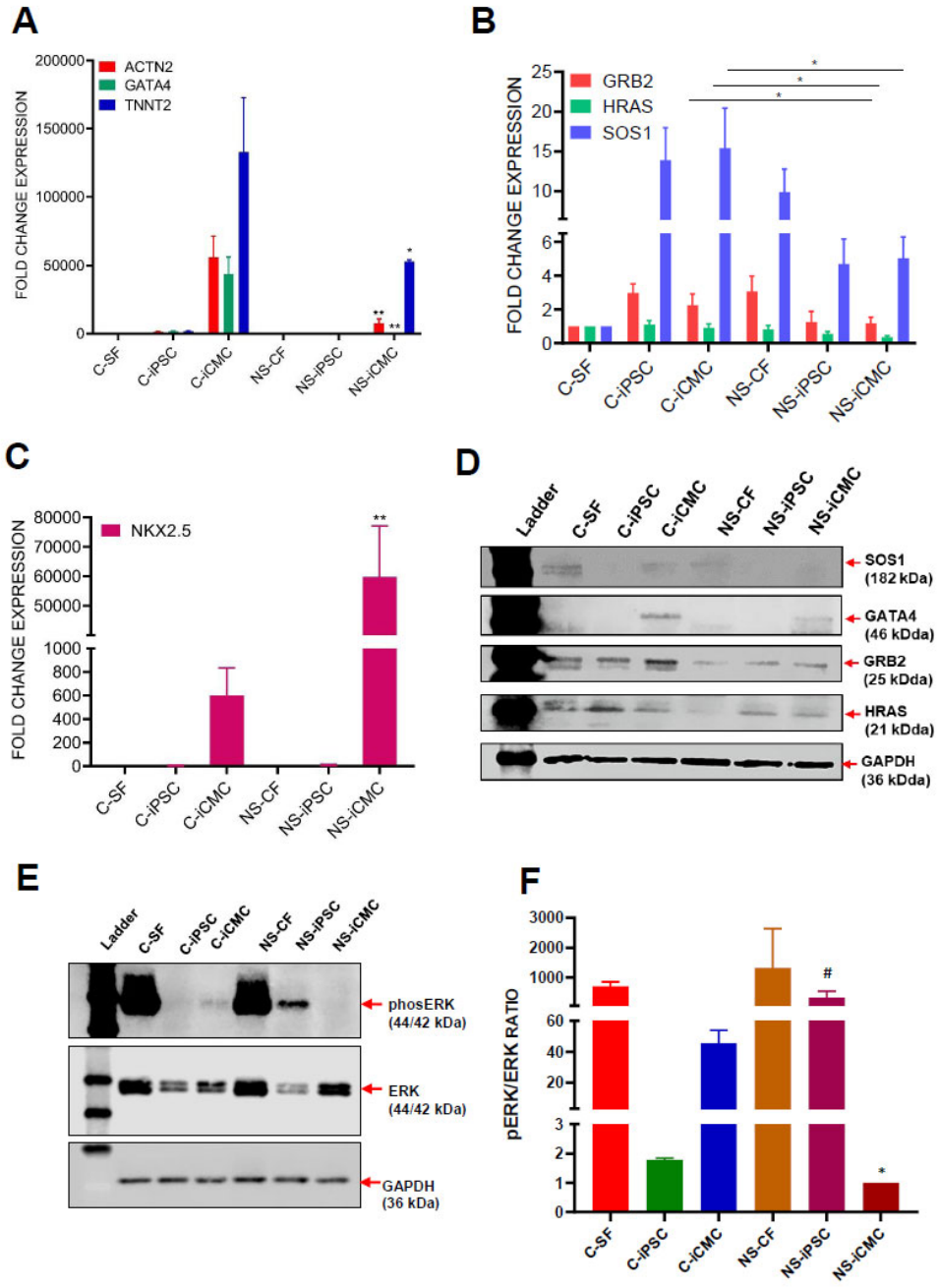

Fig. (6). *In-vitro* validation of cardiac protein expressions using NS-iCMCs in comparison with C-iCMCs. A-C, qRT-PCR analyses showing the mRNA expressions of **(A)** A decreased expression of cardiac structural genes ACTN2 and TNNT2, and a cardiac specific regulatory molecule GATA4, *P<0.05 and **P<0.01 C-iCMC *vs* NS-iCMC; **(B)** A decreased expression of *SOS1* and RAS-MAPK pathway associated genes GRB2 and HRAS *P<0.05 C-iCMC *vs* NS-iCMC; **(C)** An increased NKX2.5 gene expression observed in NS-iCMCs when compared to Noonan syndrome (NS) patient-derived cardiac fibroblasts (NS-CF), NS-iPSCs, C-iPSCs and C-iCMCs. Each bar represents the mean ± SEM of three replicated experiments. Each gene expression was normalized with 18S rRNA. *P<0.001 *vs* C-iCMC. **(D)** Western blot analyses showing the protein expressions of GATA4, GRB2, HRAS, and *SOS1* in NS-CF, NS-iPSCs and NS-iCMCs in comparison with C-SF, C-iPSCs and C-iCMC. GAPDH was used as a protein loading control. **(E)** Western immunoblotting for the phosphorylation of ERK1/2 at the Thr202/Tyr204 and the ERK1/2. GAPDH was used to identify the equal loading of protein samples. **(F)** The ratio of the phosphorylated ERK1/2 at the Thr202/Tyr204 and the total ERK1/2 was calculated. The activated ERK is significantly reduced in NS-iCMC when compared to N-iCMC *P<0.05 *vs*. NS-iCMC, whereas the activated ERK significantly increased in NS-iPSC when compared to C-iPSC $^{\#}$P<0.05 *vs*. C-iPSC. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

SOS1 promotes the formation of active RAS, leading to the activation of RAF-MEK-ERK cascade, which plays a key role in cardiac physiology/pathophysiology and cancer [36-38]. In order to analyze the effect of *SOS1* mutation in the downstream effector molecules, we studied the ERK1/2 activation by the phosphorylation of ERK1/2 at the Thr202/-Tyr204 by Western blotting. Our results indicate that ERK1/2 activation was significantly decreased in NS-iCMCs compared to C-iCMCs (Figs. **6E** and **F**). ERK1/2 were found to be activated in both control and NS parent cells; and interestingly, their activation was found to be higher in NS-iPSCs, compared to control iPSCs (Figs. **6E** and **F**). Importantly, our GAPDH loading control showed an equal loading of proteins in all the lanes whereas, the total ERK1/2 expressions were unexpectedly varied among the samples.

4. DISCUSSION

Bioinformatic tools are capable of identifying the genetic variations that are associated with a patient's disease genome. There are several bioinformatic tools available to acquire specific and accurate genetic information that is already present in online databases. Genome-wide association studies (GWAS) can be applied for testing millions of genetic variants and are also examined to identify genotype--phenotype associations across the genomes. These associations provide insights into the disease susceptibility through the identification of disease specific genes and mechanisms [39]. The SNPs in the coding region of human genes were associated with genetic disorders [14]. A large number of SNPs have been reported in the database, and it is difficult to screen all the SNPs for the particular phenotype [40, 41]. The computational analysis tools help to narrow down and examine the pathogenic SNPs for the specific genetic disorders and in minimizing the risk [41, 42].

NS is an autosomal dominant genetic disorder, where the alterations are predominantly present in four genes - *PTPN11, SOS1, RAF1*, and *KRAS*. The *SOS1* is the second dominant gene for NS [1, 2, 5]. In the human *SOS1* gene, a total of 38137 SNPs have been listed in the database. Among these, 688 nsSNPs were involved in the functional protein coding regions. In the present study, various computational tools based on different algorithms were utilized to screen the most pathogenic nsSNPs of the human *SOS1* gene to induce NS. A similar approach has been carried out to test the functional nsSNPs of *ARHGEF6, BRAF, TAGAP*, and *UTY* gene [41-44].

The 688 nsSNPs of the *SOS1* gene were clinically filtered and it was found that 16 nsSNPs were pathogenic with NS [3-13, 45-56]. Using various *in-silico* tools such as, SIFT, PolyPhen, PROVEAN, FATHMM, Mutation Taster, Condel, LRT, SNPs&GO, and PhD-SNP, we identified that 11 out of 16 nsSNPs, such as p.Met269Arg, (rs137852813), p.Arg552Gly (rs137852814), p.Arg552Ser (rs267607079), p.Trp432Arg (rs267607080), p.Glu433Lys (rs397517147), p.Gly434Arg (rs397517148), p.Ser548Arg (rs397517149), p.Ile437Thr (rs397517150), p.Leu550Pro (rs397517153), p.Arg552Thr (rs397517154), and p.Cys441Tyr (rs727504295) were more pathogenic to induce NS.

Protein stability analyses that using I-Mutant 2.0 and iPTREE-STAB revealed that all of the 11 nsSNPs had a decreased protein stability, except rs727504295. Decreased protein stability results in altered protein structure through increased aggregation, degradation, ubiquitination, and misfolding of proteins, leading to initiation of diseases [57-59]. Similarly, the prediction of the altered molecular mechanism by MutPred2 revealed a gain of intrinsic disorder in rs137852813 and rs267607080; gain of acetylation at K427 in rs267607080; altered coiled-coil in rs267607080, rs397517147, and rs397517148; gain of the helix in rs397517147 and rs397517148; altered metal binding in rs397517150 and rs727504295; loss of helix in rs397517153, rs397517154 and rs727504295; loss of proteolytic cleavage at D555 and altered transmembrane protein in rs397517153 and rs397517154; and gain of relative solvent accessibility and strand in rs727504295.

Our study results did not reveal the details of the domain or the type of mutation involved with the *SOS1* variants. A study has shown that most of the pathogenic *SOS1* mutations were clustered in pleckstrin-homology (PH), Dbl-homology (DH) or in the helical linker between PH domain and RAS exchanger motif (REM) domain. Furthermore, the *SOS1* variants p.Met269Arg, p.Arg552Gly, p.Arg552Ser, p.Ser548Arg, p.Ile437Thr, p.Leu550Pro, p.Arg552Thr, and p.Cys441Tyr were associated with Class 1 mutation, where they participated in the auto-inhibitory interaction of the DH and REM domains blocking RAS access [8]. The *SOS1* variants p.Trp432Arg, p.Glu433Lys and p.Gly434Arg were associated with Class 2 mutation, where they have been shown to modify the surface of the PH domain [8]. The 3D models were generated using I-TASSER for the wild type and mutant *SOS1* proteins, and their structural variations were analysed. Similar computational structural analysis have been carried out for various protein's models like ARHGEF6, BRAF, TAGAP, and UTY [41-44]. Besides, the HOPE tool was utilized to study the amino acid changes between the wild type and mutants. The analysis of interaction of proteins will be helpful to elucidate the network of functional proteomics [60, 61]. In this study, using STRING server, we found that *SOS1* interacts with cardiac structural and functional proteins, such as ACTN2, ACTN4, GATA4, NKX2.5, TNNI3 and TNNT2 mainly through interactions with GRB2 and HRAS. ACTN-4 is a newly discovered non-muscle alpha-actinin isoform that requires further investigation in cardiac physiology/pathophysiology [62].

Our *in-vitro* study results have demonstrated that NS-iCMCs containing *SOS1* gene variant (rs137852814) had a significantly decreased expression of *SOS1* mRNA and protein, when compared with normal C-iCMCs. NS-iCMCs also showed decreased mRNA and protein expressions of cardiac specific regulatory molecules *GATA4*, *RAS-MAPK* pathway associated *GRB2* and *HRAS,* and decreased mRNA expressions of cardiac structural genes like *ACTN2* and *TNNT2*, when compared with C-iCMCs. In contrast, the mRNA

expression level of *NKX2.5* was significantly upregulated in NS-iCMCs, when compared with C-iCMCs. The increased expression of *NKX2.5* in NS-iCMCs could be associated with cardiac hypertrophy and is in agreement with other published studies [63-65]. This increased expression of cardiac homeobox gene *NKX2.5* activated the *NPPA/ANF* and *HAND2* and resulted in the induction of cardiac hypertrophy. Similarly, in our previous study, we have found an increased expression of *NPPA* and *HAND2* in NS-iCMCs, compared to C-iCMCs [32].

SOS1 mutation has been shown to affect the stability of the protein [8], and our study results showed that *SOS1* mRNA and protein expressions were decreased in NS-iCMCs relative to the wildtype iCMCs. Our results are in agreement with the other study, where somatic mutations resulted in altered mRNA and protein expressions [66]. Studies have shown that the role of *SOS1*-mediated RAS-RAF-MEK-ERK cascade plays an important role in preserving cardiac function during physiological or pathological cardiac hypertrophy [36-38, 67, 68]. Although the role of ERK1/2 pathway during cardiac development is not fully understood, our study results indicated that ERK1/2 activation was decreased in NS-iCMCs compared to C-iCMCs, even though the NS-iCMCs had a significant amount of total ERK1/2 and it was mostly present as an inactive form. These results emphasise the importance of ERK1/2 activation in normal cardiac development. On the other hand, suppression of ERK activation facilitated the reprogramming of somatic cells [69], and accordingly, the ERK1/2 activation was reduced in normal C-iPSCs, and increased in NS-iPSCs. This results further indicated the pathophysiology associated with NS. It is also intriguing to argue whether the increased ERK1/2 activation in NS-iPSCs impairs its cardiac differentiation ability. Further studies are needed to explore the role of ERK in NS-iPSCs and NS-iCMCs during differentiation.

CONCLUSION

In conclusion, our *in-silico* analyses have identified that 11 variants of *SOS1* gene nsSNPs were more pathogenic to cause NS. Furthermore, our *in-vitro* studies using the NS patient-derived iCMCs carrying *SOS1* mutation showed a specific pattern of gene expression that results in a reduced expression of genes involved in cardiac development and an increased expression of genes involved in cardiac hypertrophy. However, further studies are required to elucidate the specific roles of *SOS1* and its interacting proteins in inducing cardiomyopathy in NS. Our current study results will be helpful for the screening of NS patients with *SOS1* mutation for the expression of more pathogenic variants. This will facilitate exploring *SOS1* interacting proteins in order to target and thereby discover a possible treatment option for NS.

LIST OF ABBREVIATIONS

SNPs	=	Single Nucleotide Polymorphisms
SOS1	=	Son Of Sevenless Homolog 1
NS	=	Noonan Syndrome
PTPN11	=	Protein Tyrosine Phosphatase Non-Receptor Type 11
RAF1	=	Raf-1 Proto-Oncogene, Serine/Threonine Kinase
KRAS	=	KRAS Proto-Oncogene, GTPase
nsSNPs	=	non-synonymous Single Nucleotide Polymorphisms
iPTREE-STAB	=	Interpretable decision tree-based method for predicting protein stability changes upon mutations
3D	=	3-Dimensional
MutPred	=	Mutation Prediction
dbSNP	=	Single Nucleotide Polymorphism Database
I-TASSER	=	Iterative Threading Assembly Refinement
RAMPAGE RNA	=	Annotation and Mapping of Promoters for Analysis of Gene Expression
GATA4	=	GATA Binding Protein 4
TNNT2	=	Troponin T2, Cardiac Type
ACTN2	=	Actinin Alpha 2
GRB2	=	Growth Factor Receptor Bound Protein 2
HRAS	=	HRas Proto-Oncogene, GTPase
iCMCs	=	induced Cardiomyocytes
NS-iCMCs	=	Noonan Syndrome patient-derived induced Cardiomyocytes
C-iCMCs	=	Normal Skin Fibroblast-derived induced Cardiomyocytes
qRT-PCR	=	Quantitative Real Time Polymerase Chain Reaction
NKX2-5	=	NK2 Homeobox 5
RAS-MAPK	=	Ras Mitogen Activated Protein Kinase
RAC1	=	Rac Family Small GTPase 1
PTM	=	Post Translational Modification
STRING	=	Protein-Protein Interaction Networks Functional Enrichment Analysis
NCBI	=	National Center for Biotechnology Information
UniProt	=	Universal Protein Resource
VEP	=	Variant Effect Predictor

dbNSFP	=	Database for nonsynonymous SNPs' functional predictions
SIFT	=	Sorting Intolerant from Tolerant
PolyPhen	=	Polymorphism Phenotyping
PROVEAN	=	Protein Variation Effect Analyser
FATHMM	=	Functional Analysis Through Hidden Markov Model
ConDel	=	Consensus Deleteriousness
LRT	=	Likelihood Ratio Test
SNPs&GO	=	Single Nucleotide Polymorphisms and Gene Ontology
PhD-SNPs	=	Predictor of human deleterious - Single Nucleotide Polymorphisms
DDG	=	delta delta Gibbs Free Energy
pH	=	Power of Hydrogen
FASTA	=	Fast Alignment
LOMETS	=	LOcal MEta-Threading-Server
ProTherm	=	Protein Thermodynamic
PDB	=	Protein Data Bank
BioLiP	=	Biologically Relevant Ligand-protein Interactions
HOPE	=	Homotopy Optimization Method for Protein Structure Prediction
TM-score	=	Template Model - Score
RMSD	=	Root Mean Square Deviation
TM-align	=	Template Model - Alignment
iPSCs	=	induced Pluripotent Stem Cells
C-SF	=	Control-Skin Fibroblast
C-iPSCs	=	Control-induced Pluripotent Stem Cells
NS-CF	=	Noonan Syndrome - Cardiac Fibroblast
NS-iPSCs	=	Noonan Syndrome - induced Pluripotent Stem Cells
cDNA	=	Complementary Deoxyribose Nucleic Acid
RNA	=	Ribose Nucleic Acid
UDG	=	uracil-DNA glycosylases
DEPC	=	Diethyl Pyrocarbonate
mRNA	=	messenger-Ribose Nucleic Acid
rRNA	=	ribosome-Ribose Nucleic Acid
GAPDH	=	Glyceraldehyde 3-phosphate dehydrogenase
ANOVA	=	Analysis of variance
C-score	=	Confidence score
GWAS	=	Genome-Wide Association Studies
PMID	=	PubMed ID
ARHGEF6	=	Rac/Cdc42 Guanine Nucleotide Exchange Factor 6
BRAF	=	B-Raf Proto-Oncogene, Serine/Threonine Kinase
TAGAP	=	T-cell activation Rho GTPase-activating Protein
UTY	=	Ubiquitously Transcribed Tetratricopeptide Repeat Containing, Y-Linked
NPPA	=	Natriuretic Peptide A
ANF	=	Atrial Natriuretic Factor
HAND2	=	Heart and Neural Crest Derivatives Expressed 2
DEL/INS	=	Deletion/Insertion
MNV	=	Multiple Nucleotide Variant
SNV	=	Single Nucleotide Variant
3`-UTR 3`	=	Untranslated Region
SEM	=	Standard Error Mean

AUTHORS' CONTRIBUTIONS

Conceptualization, J.R., V.S.; methodology, V.S., N.G., A.P., S.R.; software, V.S.; validation and formal analysis, V.S, N.G., S.R.; writing-original draft preparation, N.G., V.S.; writing-review and editing, J.R., N.G.; supervision, J.R.; funding acquisition, J.R. All authors have read and agreed to the current version of the manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data that supports the finding of the study is available within the article and its supplementary material.

FUNDING

This work was supported, in part, by the American Heart Association Grant-in-Aid 16GRNT30950010 and National Institutes of Health R01 grant R01HL141345 (to JR).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

SUPPLEMENTARY MATERIAL

Supplementary materials are available on the publisher's website.

REFERENCES

[1] Mendez, H.M.; Opitz, J.M. Noonan syndrome: A review. *Am. J. Med. Genet.,* **1985**, *21*(3), 493-506. http://dx.doi.org/10.1002/ajmg.1320210312 PMID: 3895929

[2] Nora, J.J.; Nora, A.H.; Sinha, A.K.; Spangler, R.D.; Lubs, H.A. The Ullrich-Noonan syndrome (Turner phenotype). *Am. J. Dis. Child.,* **1974**, *127*(1), 48-55. PMID: 4809794

[3] Ko, J.M.; Kim, J.M.; Kim, G.H.; Yoo, H.W. PTPN11, *SOS1*, KRAS, and RAF1 gene analysis, and genotype-phenotype correlation in Korean patients with Noonan syndrome. *J. Hum. Genet.,* **2008**, *53*(11-12), 999-1006. http://dx.doi.org/10.1007/s10038-008-0343-6 PMID: 19020799

[4] Pierpont, E.I.; Pierpont, M.E.; Mendelsohn, N.J.; Roberts, A.E.; Tworog-Dube, E.; Seidenberg, M.S. Genotype differences in cognitive functioning in Noonan syndrome. *Genes Brain Behav.,* **2009**, *8*(3), 275-282. http://dx.doi.org/10.1111/j.1601-183X.2008.00469.x PMID: 19077116

[5] Zenker, M.; Horn, D.; Wieczorek, D.; Allanson, J.; Pauli, S.; van der Burgt, I.; Doerr, H.G.; Gaspar, H.; Hofbeck, M.; Gillessen-Kaesbach, G.; Koch, A.; Meinecke, P.; Mundlos, S.; Nowka, A.; Rauch, A.; Reif, S.; von Schnakenburg, C.; Seidel, H.; Wehner, L.E.; Zweier, C.; Bauhuber, S.; Matejas, V.; Kratz, C.P.; Thomas, C.; Kutsche, K. *SOS1* is the second most common Noonan gene but plays no major role in cardio-facio-cutaneous syndrome. *J. Med. Genet.,* **2007**, *44*(10), 651-656. http://dx.doi.org/10.1136/jmg.2007.051276 PMID: 17586837

[6] Ferrero, G.B.; Baldassarre, G.; Delmonaco, A.G.; Biamino, E.; Banaudi, E.; Carta, C.; Rossi, C.; Silengo, M.C. Clinical and molecular characterization of 40 patients with Noonan syndrome. *Eur. J. Med. Genet.,* **2008**, *51*(6), 566-572. http://dx.doi.org/10.1016/j.ejmg.2008.06.011 PMID: 18678287

[7] Denayer, E.; Devriendt, K.; de Ravel, T.; Van Buggenhout, G.; Smeets, E.; Francois, I.; Sznajer, Y.; Craen, M.; Leventopoulos, G.; Mutesa, L.; Vandecasseye, W.; Massa, G.; Kayserili, H.; Sciot, R.; Fryns, J.P.; Legius, E. Tumor spectrum in children with Noonan syndrome and *SOS1* or RAF1 mutations. *Genes Chromosomes Cancer,* **2010**, *49*(3), 242-252. PMID: 19953625

[8] Lepri, F.; De Luca, A.; Stella, L.; Rossi, C.; Baldassarre, G.; Pantaleoni, F.; Cordeddu, V.; Williams, B.J.; Dentici, M.L.; Caputo, V.; Venanzi, S.; Bonaguro, M.; Kavamura, I.; Faienza, M.F.; Pilotta, A.; Stanzial, F.; Faravelli, F.; Gabrielli, O.; Marino, B.; Neri, G.; Silengo, M.C.; Ferrero, G.B.; Torrrente, I.; Selicorni, A.; Mazzanti, L.; Digilio, M.C.; Zampino, G.; Dallapiccola, B.; Gelb, B.D.; Tartaglia, M. *SOS1* mutations in Noonan syndrome: molecular spectrum, structural insights on pathogenic effects, and genotype-phenotype correlations. *Hum. Mutat.,* **2011**, *32*(7), 760-772. http://dx.doi.org/10.1002/humu.21492 PMID: 21387466

[9] Tartaglia, M.; Pennacchio, L.A.; Zhao, C.; Yadav, K.K.; Fodale, V.; Sarkozy, A.; Pandit, B.; Oishi, K.; Martinelli, S.; Schackwitz, W.; Ustaszewska, A.; Martin, J.; Bristow, J.; Carta, C.; Lepri, F.; Neri, C.; Vasta, I.; Gibson, K.; Curry, C.J.; Siguero, J.P.; Digilio, M.C.; Zampino, G.; Dallapiccola, B.; Bar-Sagi, D.; Gelb, B.D. Gain-of-function *SOS1* mutations cause a distinctive form of Noonan syndrome. *Nat. Genet.,* **2007**, *39*(1), 75-79. http://dx.doi.org/10.1038/ng1939 PMID: 17143282

[10] Roberts, A.E.; Araki, T.; Swanson, K.D.; Montgomery, K.T.; Schiripo, T.A.; Joshi, V.A.; Li, L.; Yassin, Y.; Tamburino, A.M.; Neel, B.G.; Kucherlapati, R.S. Germline gain-of-function mutations in *SOS1* cause Noonan syndrome. *Nat. Genet.,* **2007**, *39*(1), 70-74. http://dx.doi.org/10.1038/ng1926 PMID: 17143285

[11] Longoni, M.; Moncini, S.; Cisternino, M.; Morella, I.M.; Ferraiuolo, S.; Russo, S.; Mannarino, S.; Brazzelli, V.; Coi, P.; Zippel, R.; Venturin, M.; Riva, P. Noonan syndrome associated with both a new JNK-activating familial *SOS1* and a *de novo* RAF1 mutations. *Am. J. Med. Genet. A.,* **2010**, *152A*(9), 2176-2184. http://dx.doi.org/10.1002/ajmg.a.33564 PMID: 20683980

[12] Hwang, H.S.; Hwang, S.G.; Yoon, K.W.; Yoon, J.H.; Roh, K.H.; Choi, E.J. CIIA negatively regulates the Ras-Erk1/2 signaling pathway through inhibiting the Ras-specific GEF activity of *SOS1*. *J. Cell Sci.,* **2014**, *127*(Pt 8), 1640-1646. http://dx.doi.org/10.1242/jcs.139931 PMID: 24522193

[13] Findlay, G.M.; Smith, M.J.; Lanner, F.; Hsiung, M.S.; Gish, G.D.; Petsalaki, E.; Cockburn, K.; Kaneko, T.; Huang, H.; Bagshaw, R.D.; Ketela, T.; Tucholska, M.; Taylor, L.; Bowtell, D.D.; Moffat, J.; Ikura, M.; Li, S.S.; Sidhu, S.S.; Rossant, J.; Pawson, T. Interaction domains of *SOS1*/Grb2 are finely tuned for cooperative control of embryonic stem cell fate. *Cell,* **2013**, *152*(5), 1008-1020. http://dx.doi.org/10.1016/j.cell.2013.01.056 PMID: 23452850

[14] Cargill, M.; Altshuler, D.; Ireland, J.; Sklar, P.; Ardlie, K.; Patil, N.; Shaw, N.; Lane, C.R.; Lim, E.P.; Kalyanaraman, N.; Nemesh, J.; Ziaugra, L.; Friedland, L.; Rolfe, A.; Warrington, J.; Lipshutz, R.; Daley, G.Q.; Lander, E.S. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.,* **1999**, *22*(3), 231-238. http://dx.doi.org/10.1038/10290 PMID: 10391209

[15] Sachidanandam, R.; Weissman, D.; Schmidt, S.C.; Kakol, J.M.; Stein, L.D.; Marth, G.; Sherry, S.; Mullikin, J.C.; Mortimore, B.J.; Willey, D.L.; Hunt, S.E.; Cole, C.G.; Coggill, P.C.; Rice, C.M.; Ning, Z.; Rogers, J.; Bentley, D.R.; Kwok, P.Y.; Mardis, E.R.; Yeh, R.T.; Schultz, B.; Cook, L.; Davenport, R.; Dante, M.; Fulton, L.; Hillier, L.; Waterston, R.H.; McPherson, J.D.; Gilman, B.; Schaffner, S.; Van Etten, W.J.; Reich, D.; Higgins, J.; Daly, M.J.; Blumenstiel, B.; Baldwin, J.; Stange-Thomann, N.; Zody, M.C.; Linton, L.; Lander, E.S.; Altshuler, D. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature,* **2001**, *409*(6822), 928-933. http://dx.doi.org/10.1038/35057149 PMID: 11237013

[16] Chasman, D.; Adams, R.M. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. *J. Mol. Biol.,* **2001**, *307*(2), 683-706. http://dx.doi.org/10.1006/jmbi.2001.4510 PMID: 11254390

[17] Yates, C.M.; Sternberg, M.J. The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions. *J. Mol. Biol.,* **2013**, *425*(21), 3949-3963. http://dx.doi.org/10.1016/j.jmb.2013.07.012 PMID: 23867278

[18] McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl variant effect predictor. *Genome Biol.,* **2016**, *17*(1), 122. http://dx.doi.org/10.1186/s13059-016-0974-4 PMID: 27268795

[19] Capriotti, E.; Calabrese, R.; Fariselli, P.; Martelli, P.L.; Altman, R.B.; Casadio, R. WS-SNPs&GO: A web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics,* **2013**, *14*(Suppl 3), S6. http://dx.doi.org/10.1186/1471-2164-14-S3-S6

[20] Capriotti, E; Fariselli, P; Casadio, R I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure.

Nucleic acids research, **2005**, *33*(Web Server issue), W306-W310.

[21] Huang, L.T.; Gromiha, M.M.; Ho, S.Y. iPTREE-STAB: interpretable decision tree based method for predicting protein stability changes upon mutations. *Bioinformatics,* **2007**, *23*(10), 1292-1293. http://dx.doi.org/10.1093/bioinformatics/btm100 PMID: 17379687

[22] Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.A.; Lin, G.N.; Nam, H-J.; Mort, M.; Cooper, D.N.; Sebat, J.; Iakoucheva, L.M. MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. *BioRxiv,* **2017**. http://dx.doi.org/10.1101/134981

[23] Zhang, Y I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins,* **2009**, *77*(Suppl 9), 100-113. http://dx.doi.org/10.1002/prot.22588

[24] Roy, A; Yang, J; Zhang, Y COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res.,* **2012**, *40*(Web Server issue), W471-W477. http://dx.doi.org/10.1093/nar/gks372

[25] Yang, J.; Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.,* **2015**, *43*(W1), W174-81. http://dx.doi.org/10.1093/nar/gkv342 PMID: 25883148

[26] Colovos, C.; Yeates, T.O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.,* **1993**, *2*(9), 1511-1519. http://dx.doi.org/10.1002/pro.5560020916 PMID: 8401235

[27] Lovell, S.C.; Davis, I.W.; Arendall, W.B., III; de Bakker, P.I.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins,* **2003**, *50*(3), 437-450. http://dx.doi.org/10.1002/prot.10286 PMID: 12557186

[28] Venselaar, H.; Te Beek, T.A.; Kuipers, R.K.; Hekkelman, M.L.; Vriend, G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics,* **2010**, *11*, 548. http://dx.doi.org/10.1186/1471-2105-11-548 PMID: 21059217

[29] Zhang, Y.; Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins,* **2004**, *57*(4), 702-710. http://dx.doi.org/10.1002/prot.20264 PMID: 15476259

[30] Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; Jensen, L.J.; Mering, C.V. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.,* **2019**, *47*(D1), D607-D613. http://dx.doi.org/10.1093/nar/gky1131 PMID: 30476243

[31] Rajasingh, S.; Thangavel, J.; Czirok, A.; Samanta, S.; Roby, K.F.; Dawn, B.; Rajasingh, J. Generation of functional cardiomyocytes from efficiently generated human iPSCs and a novel method of measuring contractility. *PLoS One,* **2015**, *10*(8), e0134093. http://dx.doi.org/10.1371/journal.pone.0134093 PMID: 26237415

[32] Gurusamy, N.; Rajasingh, S.; Sigamani, V.; Rajasingh, R.; Isai, D.G.; Czirok, A.; Bittel, D.; Rajasingh, J. Noonan syndrome patient-specific induced cardiomyocyte model carrying *SOS1* gene variant c.1654A>G. *Exp. Cell Res.,* **2021**, *400*(1), 112508. http://dx.doi.org/10.1016/j.yexcr.2021.112508 PMID: 33549576

[33] Rajasingh, J.; Thangavel, J.; Siddiqui, M.R.; Gomes, I.; Gao, X.P.; Kishore, R.; Malik, A.B. Improvement of cardiac function in mouse myocardial infarction after transplantation of epigenetically-modified bone marrow progenitor cells. *PLoS One,* **2011**, *6*(7), e22550. http://dx.doi.org/10.1371/journal.pone.0022550 PMID: 21799893

[34] Liu, X.; Jian, X.; Boerwinkle, E. dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum. Mutat.,* **2011**, *32*(8), 894-899. http://dx.doi.org/10.1002/humu.21517 PMID: 21520341

[35] Cai, D.; Choi, P.S.; Gelbard, M.; Meyerson, M. Identification and characterization of oncogenic *SOS1* mutations in lung adenocarcinoma. *Mol. Cancer Res.,* **2019**, *17*(4), 1002-1012. http://dx.doi.org/10.1158/1541-7786.MCR-18-0316 PMID: 30635434

[36] Gallo, S.; Vitacolonna, A.; Bonzano, A.; Comoglio, P.; Crepaldi, T. ERK: A key player in the pathophysiology of cardiac hypertrophy. *Int. J. Mol. Sci.,* **2019**, *20*(9), E2164. http://dx.doi.org/10.3390/ijms20092164 PMID: 31052420

[37] Roberts, P.J.; Der, C.J. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene,* **2007**, *26*(22), 3291-3310. http://dx.doi.org/10.1038/sj.onc.1210422 PMID: 17496923

[38] Maillet, M.; van Berlo, J.H.; Molkentin, J.D. Molecular basis of physiological heart growth: Fundamental concepts and new players. *Nat. Rev. Mol. Cell Biol.,* **2013**, *14*(1), 38-48. http://dx.doi.org/10.1038/nrm3495 PMID: 23258295

[39] Tam, V.; Patel, N.; Turcotte, M.; Bossé, Y.; Paré, G.; Meyre, D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.,* **2019**, *20*(8), 467-484. http://dx.doi.org/10.1038/s41576-019-0127-1 PMID: 31068683

[40] Chen, X.; Sullivan, P.F. Single nucleotide polymorphism genotyping: Biochemistry, protocol, cost and throughput. *Pharmacogenomics J.,* **2003**, *3*(2), 77-96. http://dx.doi.org/10.1038/sj.tpj.6500167 PMID: 12746733

[41] Khimsuriya, Y.M.; Chauhan, J.B. Pathogenic predictions of non-synonymous variants and their impacts: A computational assessment of ARHGEF6 gene. *Egypt. J. Med. Hum. Genet.,* **2018**, *19*(4), 333-344. http://dx.doi.org/10.1016/j.ejmhg.2018.05.002

[42] Nailwal, M.; Chauhan, J.B. Computational analysis of high risk missense variant in human UTY gene: A candidate gene of AZFa sub-region. *J. Reprod. Infertil.,* **2017**, *18*(3), 298-306. PMID: 29062794

[43] Hussain, M.R.; Shaik, N.A.; Al-Aama, J.Y.; Asfour, H.Z.; Khan, F.S.; Masoodi, T.A.; Khan, M.A.; Shaik, N.S. *In silico* analysis of Single Nucleotide Polymorphisms (SNPs) in human BRAF gene. *Gene,* **2012**, *508*(2), 188-196. http://dx.doi.org/10.1016/j.gene.2012.07.014 PMID: 22824468

[44] Arshad, M.; Bhatti, A.; John, P. Identification and *in silico* analysis of functional SNPs of human TAGAP protein: A comprehensive study. *PLoS One,* **2018**, *13*(1), e0188143. http://dx.doi.org/10.1371/journal.pone.0188143 PMID: 29329296

[45] Brasil, A.S.; Malaquias, A.C.; Wanderley, L.T.; Kim, C.A.; Krieger, J.E.; Jorge, A.A.; Pereira, A.C.; Bertola, D.R. Co-occurring PTPN11 and *SOS1* gene mutations in Noonan syndrome: Does this predict a more severe phenotype? *Arq. Bras. Endocrinol. Metabol,* **2010**, *54*(8), 717-722. http://dx.doi.org/10.1590/S0004-27302010000800009 PMID: 21340158

[46] Neumann, T.E.; Allanson, J.; Kavamura, I.; Kerr, B.; Neri, G.; Noonan, J.; Cordeddu, V.; Gibson, K.; Tzschach, A.; Krüger, G.; Hoeltzenbein, M.; Goecke, T.O.; Kehl, H.G.; Albrecht, B.; Luczak, K.; Sasiadek, M.M.; Musante, L.; Laurie, R.; Peters, H.; Tartaglia, M.; Zenker, M.; Kalscheuer, V. Multiple giant cell lesions in patients with Noonan syndrome and cardio-facio-cutaneous syndrome. *Eur. J. Hum. Genet.,* **2009**, *17*(4), 420-425. http://dx.doi.org/10.1038/ejhg.2008.188 PMID: 18854871

[47] Slezak, R.; Luczak, K.; Kalscheuer, V.; Neumann, T.E.; Sasiadek, M.M. Noonan-like/multiple giant cell lesion syndrome in two adult patients with *SOS1* gene mutations. *Clin. Dysmorphol.,* **2010**, *19*(3), 157-160. http://dx.doi.org/10.1097/MCD.0b013e3283375886 PMID: 20305546

[48] Mascheroni, E.; Digilio, M.C.; Cortis, E.; Devito, R.; Sarkozy, A.; Capolino, R.; Dallapiccola, B.; Ugazio, A.G. Pigmented villonodular synovitis in a patient with Noonan syndrome and *SOS1* gene mutation. *Am. J. Med. Genet. A.,* **2008**, *146A*(22), 2966-2967. http://dx.doi.org/10.1002/ajmg.a.32538 PMID: 18925667

[49] Hanna, N.; Parfait, B.; Talaat, I.M.; Vidaud, M.; Elsedfy, H.H. *SOS1*: A new player in the Noonan-like/multiple giant cell lesion syndrome. *Clin. Genet.,* **2009**, *75*(6), 568-571. http://dx.doi.org/10.1111/j.1399-0004.2009.01149.x PMID: 19438935

[50] Gureasko, J.; Kuchment, O.; Makino, D.L.; Sondermann, H.; Bar-Sagi, D.; Kuriyan, J. Role of the histone domain in the autoinhibition and activation of the RAS activator Son of Sevenless. *Proc.*

Natl. Acad. Sci. USA, **2010**, *107*(8), 3430-3435.
http://dx.doi.org/10.1073/pnas.0913915107 PMID: 20133692

[51] Nyström, A.M.; Ekvall, S.; Berglund, E.; Björkqvist, M.; Braathen, G.; Duchen, K.; Enell, H.; Holmberg, E.; Holmlund, U.; Olsson-Engman, M.; Annerén, G.; Bondeson, M.L. Noonan and cardio-facio-cutaneous syndromes: Two clinically and genetically overlapping disorders. *J. Med. Genet.,* **2008**, *45*(8), 500-506.
http://dx.doi.org/10.1136/jmg.2008.057653 PMID: 18456719

[52] Digilio, M.C.; Lepri, F.; Baban, A.; Dentici, M.L.; Versacci, P.; Capolino, R.; Ferese, R.; De Luca, A.; Tartaglia, M.; Marino, B.; Dallapiccola, B. RASopathies: Clinical diagnosis in the first year of life. *Mol. Syndromol.,* **2011**, *1*(6), 282-289.
http://dx.doi.org/10.1159/000331266 PMID: 22190897

[53] Beneteau, C.; Cavé, H.; Moncla, A.; Dorison, N.; Munnich, A.; Verloes, A.; Leheup, B. *SOS1* and PTPN11 mutations in five cases of Noonan syndrome with multiple giant cell lesions. *Eur. J. Hum. Genet.,* **2009**, *17*(10), 1216-1221.
http://dx.doi.org/10.1038/ejhg.2009.44 PMID: 19352411

[54] Jongmans, M.C.; Hoogerbrugge, P.M.; Hilkens, L.; Flucke, U.; van der Burgt, I.; Noordam, K.; Ruiterkamp-Versteeg, M.; Yntema, H.G.; Nillesen, W.M.; Ligtenberg, M.J.; van Kessel, A.G.; Kuiper, R.P.; Hoogerbrugge, N. Noonan syndrome, the *SOS1* gene and embryonal rhabdomyosarcoma. *Genes Chromosomes Cancer,* **2010**, *49*(7), 635-641.
http://dx.doi.org/10.1002/gcc.20773 PMID: 20461756

[55] Chen, P.C.; Wakimoto, H.; Conner, D.; Araki, T.; Yuan, T.; Roberts, A.; Seidman, C.; Bronson, R.; Neel, B.; Seidman, J.G.; Kucherlapati, R. Activation of multiple signaling pathways causes developmental defects in mice with a Noonan syndrome-associated *SOS1* mutation. *J. Clin. Invest.,* **2010**, *120*(12), 4353-4365.
http://dx.doi.org/10.1172/JCI43910 PMID: 21041952

[56] Narumi, Y.; Aoki, Y.; Niihori, T.; Sakurai, M.; Cavé, H.; Verloes, A.; Nishio, K.; Ohashi, H.; Kurosawa, K.; Okamoto, N.; Kawame, H.; Mizuno, S.; Kondoh, T.; Addor, M.C.; Coeslier-Dieux, A.; Vincent-Delorme, C.; Tabayashi, K.; Aoki, M.; Kobayashi, T.; Guliyeva, A.; Kure, S.; Matsubara, Y. Clinical manifestations in patients with *SOS1* mutations range from Noonan syndrome to CFC syndrome. *J. Hum. Genet.,* **2008**, *53*(9), 834-841.
http://dx.doi.org/10.1007/s10038-008-0320-0 PMID: 18651097

[57] Spear, E.D.; Hsu, E.T.; Nie, L.; Carpenter, E.P.; Hrycyna, C.A.; Michaelis, S. *ZMPSTE24* missense mutations that cause progeroid diseases decrease prelamin A cleavage activity and/or protein stability. *Dis. Model. Mech.,* **2018**, *11*(7), dmm033670.
http://dx.doi.org/10.1242/dmm.033670 PMID: 29794150

[58] Clausen, L.; Abildgaard, A.B.; Gersing, S.K.; Stein, A.; Lindorff-Larsen, K.; Hartmann-Petersen, R. Protein stability and degradation in health and disease. *Adv. Protein Chem. Struct. Biol.,* **2019**, *114*, 61-83.
http://dx.doi.org/10.1016/bs.apcsb.2018.09.002 PMID: 30635086

[59] Plemper, R.K.; Hammond, A.L. Protein degradation in human disease. *Prog. Mol. Subcell. Biol.,* **2002**, *29*, 61-84.
http://dx.doi.org/10.1007/978-3-642-56373-7_5 PMID: 11908073

[60] Safaei, A.; Rezaei Tavirani, M.; Arefi Oskouei, A.; Zamanian Azodi, M.; Mohebbi, S.R.; Nikzamir, A.R. Protein-protein interaction network analysis of cirrhosis liver disease. *Gastroenterol. Hepatol. Bed Bench,* **2016**, *9*(2), 114-123.
PMID: 27099671

[61] Chen, S.J.; Liao, D.L.; Chen, C.H.; Wang, T.Y.; Chen, K.C. Construction and analysis of protein-protein interaction network of heroin use disorder. *Sci. Rep.,* **2019**, *9*(1), 4980.
http://dx.doi.org/10.1038/s41598-019-41552-z PMID: 30899073

[62] Hein, S.; Block, T.; Zimmermann, R.; Kostin, S.; Scheffold, T.; Kubin, T.; Klövekorn, W.P.; Schaper, J. Deposition of nonsarcomeric alpha-actinin in cardiomyocytes from patients with dilated cardiomyopathy or chronic pressure overload. *Exp. Clin. Cardiol.,* **2009**, *14*(3), e68-e75.
PMID: 20098571

[63] Thompson, J.T.; Rackley, M.S.; O'Brien, T.X. Upregulation of the cardiac homeobox gene Nkx2-5 (CSX) in feline right ventricular pressure overload. *Am. J. Physiol.,* **1998**, *274*(5), H1569-H1573.
PMID: 9612365

[64] Saadane, N.; Alpert, L.; Chalifour, L.E. Expression of immediate early genes, GATA-4, and Nkx-2.5 in adrenergic-induced cardiac hypertrophy and during regression in adult mice. *Br. J. Pharmacol.,* **1999**, *127*(5), 1165-1176.
http://dx.doi.org/10.1038/sj.bjp.0702676 PMID: 10455263

[65] Bär, H.; Kreuzer, J.; Cojoc, A.; Jahn, L. Upregulation of embryonic transcription factors in right ventricular hypertrophy. *Basic Res. Cardiol.,* **2003**, *98*(5), 285-294.
http://dx.doi.org/10.1007/s00395-003-0410-2 PMID: 12955401

[66] Jia, P.; Zhao, Z. Impacts of somatic mutations on gene expression: An association perspective. *Brief. Bioinform.,* **2017**, *18*(3), 413-425.
PMID: 27127206

[67] Mutlak, M.; Schlesinger-Laufer, M.; Haas, T.; Shofti, R.; Ballan, N.; Lewis, Y.E.; Zuler, M.; Zohar, Y.; Caspi, L.H.; Kehat, I. Extracellular signal-regulated kinase (ERK) activation preserves cardiac function in pressure overload induced hypertrophy. *Int. J. Cardiol.,* **2018**, *270*, 204-213.
http://dx.doi.org/10.1016/j.ijcard.2018.05.068 PMID: 29857938

[68] Chen, P.C.; Yin, J.; Yu, H.W.; Yuan, T.; Fernandez, M.; Yung, C.K.; Trinh, Q.M.; Peltekova, V.D.; Reid, J.G.; Tworog-Dube, E.; Morgan, M.B.; Muzny, D.M.; Stein, L.; McPherson, J.D.; Roberts, A.E.; Gibbs, R.A.; Neel, B.G.; Kucherlapati, R. Next-generation sequencing identifies rare variants associated with Noonan syndrome. *Proc. Natl. Acad. Sci. USA,* **2014**, *111*(31), 11473-11478.
http://dx.doi.org/10.1073/pnas.1324128111 PMID: 25049390

[69] Huh, S.; Song, H.R.; Jeong, G.R.; Jang, H.; Seo, N.H.; Lee, J.H.; Yi, J.Y.; Lee, B.; Choi, H.W.; Do, J.T.; Kim, J.S.; Lee, S.H.; Jung, J.W.; Lee, T.; Shim, J.; Han, M.K.; Lee, T.H. Suppression of the ERK-SRF axis facilitates somatic cell reprogramming. *Exp. Mol. Med.,* **2018**, *50*(2), e448.
http://dx.doi.org/10.1038/emm.2017.279 PMID: 29472703