© IJHG 2020 Int J Hum Genet, 20(2): 55-65 (2020)
PRINT: ISSN 0972-3757 ONLINE: ISSN 2456-6330 DOI: 10.31901/24566330.2020/20.02.738

Nuclear Co-repressor 1: A Potential Candidate Gene in the Manifestation of Congenital Heart Diseases

Rahim Gholipoorfeshkecheh[1], Swati Agarwala[1], Santhosh Krishnappa[2], Mysore R. Savitha[3], Narayanappa Doddaiah[4] and Nallur B. Ramachandra[1*]

[1]Department of Studies in Genetics and Genomics, University of Mysore, Mysuru, India
[2]Sri Jayadeva Institute of Cardiovascular Science & Research, Mysuru, India
[3]Mysore Medical College & Research Institute, Mysuru, India
[4]JSS Medical College, JSS Academy of Higher Education and Research, Mysuru, India

KEYWORDS Histone Deacetylase. Non-syndromic. Non-synonymous. Single Nucleotide Polymorphism. Whole Exome Sequencing. Ventricular Septal Defects

ABSTRACT The heart is the first organ formed during embryogenesis. The development of the heart is controlled by a network of pathways involving various transcription factors and signalling. Disruptions of these factors lead to the manifestation of congenital heart disease (CHD). The ventricular septal defects (VSD) of CHD are more frequent in the Indian population, particularly in Mysuru. In view of this, to identify defective genes in ventricular septal development, whole-exome sequencing was performed in seven cases of non-syndromic VSDs and three cases of Tetralogy of Fallot. The exome data analysis revealed that of all the defective genes identified, *Nuclear Co-repressor* 1 (*NCoR1*) is the prominent gene. It showed four non-synonymous damaging single nucleotide variants, that is, G244A, C241T, G207C and G121A. These mutations were present in the GPS2 protein domain, an integral part of NCoR1-HDAC complex involved in myogenesis. Further, pathway enrichment and network analysis indicated that NCoR1 interacts with HEY1, HEY2, MEF2A, NOTCH2 and HDAC3 proteins involved in heart development. Hence, defects in NCoR1 are responsible for VSD in heart development.

INTRODUCTION

The heart is the first organ formed during embryogenesis. Malformation of the heart or the blood vessels associated leads to congenital heart diseases (CHDs) affecting various parts or functions. CHDs are amongst the primary complex diseases of interest for human health, especially in India (Smitha et al. 2006). The prevalence of CHDs in India is 8.07 per 1000 live births (Saxena et al. 2016) and is 8.0 per 1000 live births at the global level, but this figure varies worldwide (Chelo et al. 2016). The difference in prevalence is due to the age at diagnosis of CHD and the screening approach.

Various recognisable genetic factors involved in the manifestation of CHDs have been identified. Several linkage studies have identified DNA variants relevant to CHDs as a leading cause in its manifestation (Okubo et al. 2004). Genome-wide linkage analysis of CHD cases identified genes, Selenoprotein S (*SELS*), Small Nuclear Ribonucleoprotein Polypeptide A (*SNRPA*1) and Proprotein Convertase Subtilisin/Kexin Type 6 (*PCSK*6) on 15q26.3 and Transcription Factor 4 (*TCF*4) on 18q21.2 (Flaquer et al. 2013). Earlier studies from the laboratory reported two Single Nucleotide Polymorphisms (SNPs) of NK2 Homeobox 5 (*NKX2.5*) and five non-synonymous SNPs of GATA Binding Protein 4 (*GATA4*) in different CHD cases (Dinesh et al. 2010). The researchers also reported that SNPs in Cysteine Rich with EGF Like Domain 1 (*CRELD*1) change amino acids arginine to cysteine in the second calcium-binding epidermal growth factor (EGF) domain, leading to change in the β-sheet in the secondary structure in CHD cases (Kusuma et al. 2011). Several studies have reported that about 400 genes are associated with CHDs, and among them, about 40 are causal genes (Cao et al. 2016; Blue et al. 2017). Recently, Peyvandi et al. (2019) have reported that infants with severe CHDs showed abnormalities in brain maturity, indicating the risk for developing brain injuries.

Cardiac septation defects are the most common form of CHD and account for approximately fifty percent of all CHDs. Amongst all the various categories of CHDs, ventricular septal defects (VSDs) are the most common (15% to 20%) with a

[*]Address for correspondence:
Nallur B. Ramachandra Ph. D, FASc
Phone: +91-0821- 2419888
E-mail: ramachandra@genetics.uni-mysore.ac.in

prevalence rate of 2.62 per 1000 live births worldwide (Menting et al. 2015). It can result from the incomplete proliferation of cardiomyocytes in the muscular septum (Rosenfeld et al. 2006). In India, the prevalence of VSD is the highest (33%) amongst all the forms of CHDs (Bhardwaj et al. 2015), and in Mysore, it is at 40.97 percent (Smitha et al. 2006). Some of the prominent genes involved in VSDs are *GATA4*, *NKX2.5*, T-Box Transcription Factor 5 (*TBX5*), *TBX20* and *TBX1* as well as for Tetralogy of Fallots (TOFs) are *NKX2.5*, Jagged Canonical Notch Ligand 1 (*JAG1*), *TBX1*, Notch Receptor 1 (*NOTCH1*), and *NOTCH2* (Dinesh et al. 2010; Kusuma et al. 2011; Muntean et al. 2017). SNP arrays and genome-wide association studies (GWAS) have been used to detect abnormalities in the prenatal diagnosis of VSDs (Fu et al. 2017). Multiple such studies have identified SNPs and copy number abnormalities playing a role in the pathogenesis of CHDs. Further, missense variations of *TBX5* in cases with isolated VSDs were identified by Sanger sequencing (Chen et al. 2017).

Many of these studies have focused on common single nucleotide variants (SNVs) without considering rare variants. Further, exome sequencing harbours about eighty-five percent of mutations inclusion of rare exonic variants with significant effects on disease-related traits (Majewski et al. 2011). Therefore, whole-exome sequence analysis acts as a reliable tool to identify SNVs in CHDs.

Objectives

The present study aimed at identifying damaging and high-risk DNA variants in non-syndromic VSD and TOF cases using whole-exome sequence dataset. Further, to identify how these gene-gene interactions affect the complete or partial cascade of a downstream process leading to CHD manifestation and establish a gene-gene interaction network for these interactors.

MATERIAL AND METHODS

Samples

In the present study, ten severe cases of non-syndromic CHD were subjected to whole-exome sequence based on inclusion-exclusion criteria from Cheluvamba hospital, Jayadeva hospital and JSS hospitals in Mysuru.

Diagnostic Criteria

An echocardiogram test, defect and size, misplaced aorta, hypertrophy and valve stenosis were considered for the selection of VSD and TOF subjects. Karyotyping was performed to filter out any trisomy or chromosomal anomaly cases.

Inclusion Criteria

In this study, clinically diagnosed, non-syndromic CHD cases with VSDs and TOFs having the defect and size above 5mm were recruited.

Exclusion Criteria

Subjects with chromosomal abnormalities and other CHD malformations like Atrial Septal Defect (ASD), Patent Ductus Arteriosus (PDA), Aortic Stenosis (AS), Pulmonary Stenosis (PS), Atrioventricular Septal Defect (AVSD), Pulmonary Artesia (PA) and Tricuspid Artesia (TA) were excluded.

Ethics Statement

The ethical approval was acquired from the Institutional Human Ethics Committee (IHEC-UOM No138/PhD/2016-17) of University of Mysore, IHEC of Mysore Medical College (MMC EC 84/16) and IHEC of JSS Medical College, Mysore (JSSMC/IEC/33/1570/2017/18) for this study. Informed consent was obtained from all the subjects enrolled in this research.

DNA Isolation

Genomic DNA was extracted from 2.5 ml human peripheral whole blood collected through vacutainer collection tubes using the Qiagen kit (Cat. No. 51104) following the manufacturer's standard protocol. The isolated DNA was stored at -20 °C for WES.

Whole Exome Sequencing and Variant Calling

A total of seven VSDs and three TOFs subjects were selected for WES to identify causal

genes and novel variants. The pooled library was prepared by using the TruSeq Library Preparation kit from all 10 Quality Check (QC) qualified samples. QC of the library was performed using a 4200 Tape station system (Agilent Technologies). The pooled library was sequenced on NextSeq 500 using 2×75 bp chemistry at the coverage of 100X. WES analysis was performed using standard Illumina protocols. The raw data were obtained in FASTQ format, and the high-quality reads were aligned against the human reference genome (hg19) and mapped on the reference genome using Burrows-Wheeler Aligner (BWA) build of the human reference genome. Post alignment quality check was performed, and variants that failed QC were removed.

Variant Annotation and Variant Function Prediction

Variant annotation was performed by using various software such as wANNOVAR (http://wannovar.wglab.org/), based on various algorithms with statistical significance. Damaging non-synonymous variants was predicted by sorting intolerant from tolerant (SIFT) (http://sift.bii.a-star.edu.sg/), polymorphism phenotyping V2 (PolyPhen-2) (http://genetics.bwh.harvard.edu/pph2/) and scoring and mutation taster (http://www.mutationtaster.org/). Damaging non-synonymous variants was considered as a protein-altering substitution. Finally, these damaging non-synonymous variants were interrogated in candidate genes in study subjects.

Network Analysis

Network analysis was performed on software tools like GENEMANIA (https://genemania.org/) and STRING (https://string-db.org/), which are based on exhaustive curated literature. Haplo-insufficiency score and Probability Loss Intolerance (PLI) values were derived from the Exome Aggregation Consortium (ExAC) browser to identify the effect of gene variants in their protein products.

RESULTS

The clinical features of all the ten CHD cases are as follows. All the seven VSD cases showed several episodes of dilations. Of the seven VSD cases, two cases had pulmonary artery dilation, two cases had right ventricular dilations, and three cases had left ventricular dilations. All the three TOF cases showed sub-aortic right ventricular hypertrophy, pulmonary valve stenosis and defects in ventricular septa and misplaced aorta through an echocardiogram. All the TOF cases showed aortic overriding and right ventricular hypertrophy. Of the three TOF cases, two subjects had undergone cardiac surgery.

Whole exome sequence analysis resulted in an average coverage of 95.5 percent with a mean of 5.4 bases at 100X coverage. Overall alignment against the reference human genome hg19 was estimated at 87.5 percent. With three million reads, 78,405 variants were identified with confidence variant filtration. By performing multiple layers of gene variant filtration, 24 genes containing 41 damaging non-synonymous SNVs and frame shift deletions were identified from whole-exome sequence dataset (Table 1). Further, based on filtration with respect to allele frequency and pathogenicity, ten variants in six genes, namely, Mastermind Like Transcriptional Coactivator 3 (*MAML*3), Hairy/Enhancer-of-split related with YRPW motif protein 2 (*HEY*2), Myocyte Enhancer Factor 2A (*MEF*2*A*), *NCoR*1, *NOTCH*2 and Histone Deacetylase 9 (*HDAC*9) were observed (Table 2, Fig. 1). *NCoR*1 harboured four variants, G244A, C241T, G207C, and G121A on exon 2 in the transcript NM_001190438. All the ten subjects had the variants, rs76780359, rs78230791 and rs200020868, as well as the eight subjects, had the variant, rs74453660. All these variants were present in the protein domain G protein suppressor 2 (GPS2) (Fig. 1) with multiple physical interactors.

The NCoR1 protein interacts mainly with Recombination Signal Binding Protein for Immunoglobulin Kappa J Region (RBPJ), HDCA9, HDCA3 and HEY2 proteins, which are involved in heart development. These physical interactors also showed SNVs present across other samples, affecting the downstream cascade of pathways and processes resulting in disease manifestation. One major pathway affected by NCoR1 and its physical interactors was the histone deacetylase machinery (Fig. 2).

Table 1: Summary of damaging gene variations found in exome sequencing data of 10 samples

S. No.	Gene	rsID	Ref/Alt	Mutation type	TOF-1	TOF-2	TOF-3	VSD-1	VSD-2	VSD-3	VSD-4	VSD-5	VSD-6	VSD-7	%
1	*NCoR1*	rs76780359	G/A	Nonsynonymous SNV	+	+	+	+	+	+	+	+	+	+	100
		rs78230791	C/T	Stopgain	+	+	+	+	+	+	+	+	+	+	100
		rs200020868	G/C	Nonsynonymous SNV	+	+	+	+	+	+	+	+	+	+	100
2	*ACTN2*	rs869075326	C/-	Frameshift deletion	+	+	+	+	+	+	+	+	-	+	90
3	*NCoR1*	rs74453660	G/A	Nonsynonymous SNV	+	+	+	+	+	-	+	+	+	-	80
4	*MAP2K3*	rs74575904	T/G	Nonsynonymous SNV	+	+	+	+	+	+	+	+	+	+	70
5	*NOTCH2*	rs372504208	GG/-	Frameshift deletion	+	+	+	-	-	-	+	+	-	+	60
6	*NFATC1*	rs754093	T/G	Nonsynonymous SNV	-	-	+	-	+	+	+	+	+	-	60
7	*HEY1*	rs142613628	GGCA TGT/-	Frameshift deletion	-	+	-	-	-	+	-	+	+	+	50
8	*TBX1*	rs172562	C/T	Nonsynonymous SNV	+	-	-	+	-	+	-	+	-	+	50
9	*MAML3*	rs5862430	TGCTGC	Frameshift deletion	+	-	+	-	+	-	+	-	+	-	50
10	*IRX2*	rs76906087	C/A	Nonsynonymous SNV	-	-	-	+	+	-	+	+	+	-	50
11	*LFNG*	rs764380644	GATG/-	Frameshift deletion	+	+	-	+	-	+	-	-	-	-	40
12	*MTHFR*	rs1801133	G/A	Nonsynonymous SNV	+	-	-	+	-	+	-	+	-	-	40
13	*FOXC1*	rs572346201	-/CGG	Nonframeshift insertion	-	-	+	+	-	-	+	+	-	-	40
14	*MEF2A*	rs58424802	CAG CAG/-	Nonframeshift deletion	+	+	-	-	+	+	-	-	-	-	40
15	*JAG1*	rs35761929	G/C	Nonsynonymous SNV	+	-	+	-	-	-	-	+	-	-	30
16	*MYH7B*	rs11906160	G/A	Nonsynonymous SNV	-	-	+	+	-	+	-	-	-	-	30
17	*NPPA*	rs5063	C/T	Nonsynonymous SNV	-	-	-	+	-	-	-	-	+	+	30
18	*HEY2*	rs549151246	A/G	Nonsynonymous SNV	-	-	-	-	-	+	-	+	-	-	20
19	*LFNG*	rs778766046	-/GATG	Frameshift insertion	-	-	-	-	-	-	+	-	+	-	20
20	*MYH6*	rs28711516	C/T	nonsynonymous SNV	+	-	-	+	-	-	-	-	-	-	20
21	*CHD7*	rs370129047	A/G	Nonsynonymous SNV	+	+	-	-	-	-	-	-	-	-	20
22	*NPPA*	rs5065	A/G	Stoploss	+	-	-	-	-	-	+	-	-	-	20
23	*ZFPM1*	rs71395304	G/T	Nonsynonymous SNV	-	-	+	-	-	-	+	-	-	-	20
24	*FOXC1*	rs778279987	-/GGC	Nonframeshift insertion	-	+	-	-	-	-	-	-	-	-	10
			GCG/-	nonframeshift deletion	-	-	-	-	-	-	-	-	-	+	10
			CGG/-	Nonframeshift deletion	-	-	-	-	+	-	-	-	-	-	10
25	*MYH11*	rs587781055	A/G	Nonsynonymous SNV	-	-	-	+	-	-	-	-	-	-	10
26	*GATA4*	rs115099192	C/A	Nonsynonymous SNV	-	-	-	-	-	+	-	-	-	-	10
27	*MAML3*	rs5862430	CTG/-	Nonframeshift deletion	-	-	-	+	-	-	-	-	-	-	10
28	*HEY1*	rs752681708	C/T	Nonsynonymous SNV	-	-	-	-	-	-	+	-	-	-	10
29	*MYH6*	rs28711516	C/G	Nonsynonymous SNV	-	+	-	-	-	-	-	-	-	-	10
30	*HDAC9*	rs542095233	G/C	Nonsynonymous SNV	-	-	+	-	-	-	-	-	-	-	10
31	*PTK2B*	No rsID	G/T	Nonsynonymous SNV	-	-	-	-	-	-	+	-	-	-	10

Table 1: Contd...

S. No.	Gene	rsID	Ref/Alt	Mutation type	TOF-1	TOF-2	TOF-3	VSD-1	VSD-2	VSD-3	VSD-4	VSD-5	VSD-6	VSD-7	%
		8:27279825-27279825													
32	*MEF2A*	No rsID	CAG/-	Nonframeshift deletion	+	-	-	-	-	-	-	-	-	-	10
		15:100252780-100252781	AA/-	Frameshift deletion	-	-	+	-	-	-	-	-	-	-	10
			AG/-	Frameshift deletion	-	-	+	-	-	-	-	-	-	-	10
			CAGCA	Nonframeshift deletion	-	-	-	-	-	-	-	+	-	-	10
			GCAG/-CAG	Nonframeshift deletion	-	-	-	-	-	-	-	+	-	-	10
			CAG/-CAG/-	Nonframeshift deletion	-	-	-	-	-	-	-	-	-	+	10

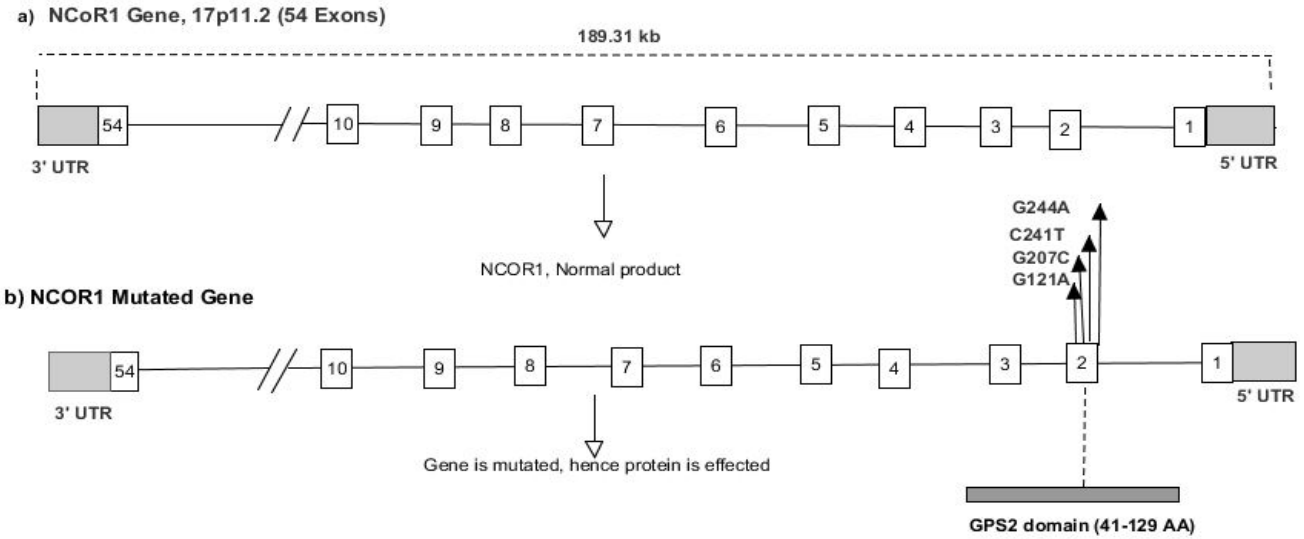

Fig. 1. a) Characterization of *NCoR1*: a) Structural Organization of *NCoR1*. The *NCoR1* positioned on 17p11.2 with 54 exons, covered 189.31 kb, from 16121499 to 15932191 (NCBI 37) on the reverse strand. b) Structural organization of *NCoR1* in cases with four nonsynonymous damaging SNVs: G121A, G207C, C241T and G244A. All these variants were present in the protein domain G protein suppressor 2 (GPS2) domain, an integral part of NCoR1-HDAC complex involved in myogenesis

Table 2: Damaging nonsynonymous heterozygous gene variants identified in non-syndromic VSD and TOF in different transcripts of Whole-exome sequence data

Gene	*rsID*	*Variant position*	*No. of samples*	*Types of change*	*Haplo-insufficiency score (in %)**	*PLI***
NCOR1	rs76780359	17:16068340-16068340	10	exon2:c.G244A:p.E82K	2.41	1
	rs78230791	17:16068343-16068343	10	exon2:c.C241T:p.R81X		
	rs200020868	17:16068377-16068377	10	exon2:c.G207C:p.K69N		
	rs74453660	17:16068463-16068463	8	exon2:c.G121A:p.G41R		
NOTCH2	rs372504208	1:120612003-120612004	6	exon1:c.17_18del:p.P6R	3.2	1
MAML3	rs5862430	4:140811064-140811069	5	exon3:c.1513_1514del:p.Q505T	18.47	0.33
	rs5862430	4:140651585-140651587	1	exon4:c.2302_2304del:p.768_768del		
MEF2A	rs58424802	15:100252710-100252715	4	exon8:c.1024_1029del:p.342_343del	6.03	0.74
HEY2	rs549151246	6:126080304-126080304	2	exon5:c.A370G:p.I124V	9.6	0.44
HDAC9	rs542095233	7:19015474-19015474	1	exon24:c.G2936C:p.C979	0.05	1

***High ranks (0-10 %) indicate a gene exhibit more haplo-insufficiency**
**** Probability Loss Intolerance (PLI) from the Exome Aggregation Consortium (ExAC) browser, (>0.9) indicate extremely Loss of Function (LOF), genes with low PLI scores (PLI ≤ 0.1) are LOF tolerant**

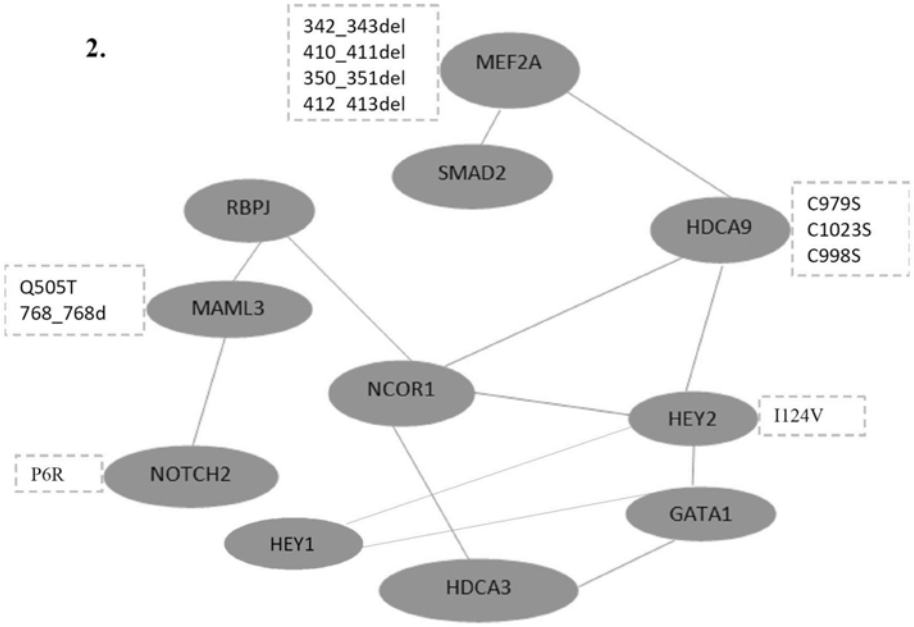

Fig. 2. Physical interaction of NCoR1 with RBPJ, HDCA9, HDCA3 and HEY2 proteins: the players in CHD manifestation.
These gene-gene interactions showed damaging mutations present among various samples, affecting the downstream cascade of pathways and processes resulting in disease manifestation. One major pathway affected by NCoR1 and its physical interactors was the histone deacetylase machinery

DISCUSSION

Proper heart development is crucial for its function. Several proteins and their timely interaction in early embryogenesis are important for the development of a normal heart. If the genes coding for these proteins is defective, then an abnormal heart is formed. Unravelling the defective genes in various types of CHDs is required to understand the causative factors of the disease and possible diagnosis and treatment. One such attempt has been made in the present study by whole-exome sequencing the severe cases with CHDs. The prominent cases of CHDs in India are VSDs and TOFs (Smitha et al. 2006).

Severe complications occur in VSDs, including severe pulmonary hypertension, aortic regurgitation, aortic cusp prolapse, left ventricle and right ventricle outflow tract obstruction (Chaudhry et al. 2011). Echocardiogram has shown several complications in VSD and TOF cases. The cardiologist offered some of these cases only medical management. Four cases of VSD were appropriate for percutaneous device closure. One of the VSD cases underwent device closure successfully. Two of the VSD cases were in an inoperable condition, medically called an Eisenmenger phenomenon. Out of three TOF cases, two cases underwent intracardiac repair, and one was managed medically with antibiotics because of endocarditis.

Whole exome sequence analysis filtered all the variations and identified defects in six genes, namely, *MAML*3, *HEY*2, *MEF*2*A*, *NCoR*1, *NOTCH*2 and *HDAC*9. All these genes, except *MAML*3, showed a less than ten percent haploinsufficiency score indicating Loss of Function (LOF) variants in one gene copy leading to the reduced amount of protein product. NCoR1, NOTCH2 and HDAC9 represented the extreme loss of function variants. The PLI score explains the probability of genes that belonged in the hap-

lo-insufficient group, resulting in extremely intolerant loss of function variants. Based on the present investigation, amongst all these genes identified, *NCoR*1 with several damaging, nonsynonymous variants with compound heterogeneity and allele frequency less than 0.5, expressed in the left ventricle stands out and hence studied in detail. Further, the role of *NCoR*1 studied by various investigators using different techniques and model organisms from 1999 to 2019 has been summarised in detail in Supplementary Table 1.

*NCoR*1 is mapped at 17p11.2 with 189.31 kb, from 16121499 to 15932191 (NCBI 37) on the reverse strand. *NCoR*1 contains several conserved functional domains with few intrinsically folded regions (Wen et al. 2000). *NCoR1* is primarily expressed in the left ventricle of the heart. The four variants of *NCoR*1 identified in the present study map to the GPS2 protein domain. Variants in the GPS2 protein domain render the domain non-functional, failing downstream processes. Besides, the NCoR/HDAC3 complex specifically functions by blocking myoblast determination protein 1 (MYOD1) in myogenesis. The knockout of the NCoR/HDAC3 complex demonstrated the disruption of blocking differentiation factors (Phelps et al. 2016). Therefore, NCoR is necessary for repression of myogenesis by the effect on MYOD. NCoR regulates and controls the activity of MYOD in mammalian differentiation. MYOD has a role during myogenesis, activating both muscle-specific gene transcription and regulation of cell cycle (Bailey et al. 1999). *NCoR* is also needed to limit fibroblast growth factor 8 (*FGF8*) expression in caudal progenitors and heart progenitors (Kumar et al. 2016).

NCoR1 is a transcriptional co-regulator that controls the activity of many transcriptional factors such as myocyte enhancer factor 2 (*MEF2*) involved in heart development. The receptor interaction domains of NCoR1 interacted with MEF2a to repress its transcriptional activity and in turn, formed a complex with MEF2a and class IIa Histone deacetylases (*HDACs*) to suppress hypertrophy related genes. Further, NCoR1, along with intercellular adhesion molecule 1 (*ICAM1)* and AKT Serine/Threonine Kinase 3 (*AKT3*) has been important in cardiogenesis of embryos (Lin et al. 2018; Li et al. 2019) reported that the knockdown and overexpression of *NCoR1* in mice model system resulted in exacerbation and mitigation, respectively in cardiomyocyte hypertrophy- a major characteristic of Tetralogy of Fallot.

All the variants of *NCoR*1 were observed within the GPS2 domain, which affects the NCoR1-HDAC3 complex. Disruption of histone deacetylation results in impaired transcription regulation of myoblast differentiation, in turn affect heart development. Therefore, NCoR1 is a key component of the co-repressor complex that plays an important role in the control of heart development and differentiation.

The development of the heart is controlled by a network of transcription factors and pathways of gene signalling involved in heart development and differentiation. Any change in these cardiac genes and signalling pathways could cause CHDs. Further, NCoR1 protein interacts with RBPJ, HDAC9, HDAC3 and HEY2 proteins, which are expressed in the left ventricle, artery and aorta. RBPJ and HEY2 are downstream effectors of the NOTCH signalling pathway required for heart development. HDACs also involved in the development and cell cycle process, so any disruption in this interaction, triggers CHD manifestation. The HDACs eliminate the acetyl group from the DNA histone leading to the coiling of chromatin, which inhibits transcription activity (Meyners et al. 2017). DNA histone modification has involved in cardiac fibroblast activation (Tao et al. 2018). Therefore, the corepression complex of NCoR1-HDAC and other physical interactors disrupted by the *NCoR1* mutations leads to impaired heart development resulting in CHD manifestation.

CONCLUSION

Based on the current investigation, *NCoR*1 is the potential candidate gene involved in heart development accounting for the manifestation of VSD and TOF. Identification and occurrence of *NCoR1* SNVs among all the ten samples with haplo-insufficiency of 2.41 in exon 2 indicated that it is a potential candidate gene for CHD in general and VSD and TOF in particular. This study would serve as a pilot attempt and frontier for identification of causal genes and pathways impaired for VSD and TOF manifestation exclusively in the Indian population. Further, this data will be validated in larger Indian samples.

Supplementary Table 1: Summary of studies on *NCoR1* in morphogenesis from 1999 to 2019

S. No.	*Type of study*	*Sample*	*Key findings*	*Phenotype*	*Reference*
1	Gene expression	-	NCoR regulate the activity of MYOD and myogenesis	-	Bailey et al. 1999
2	Affinity purification approach	HeLa cell extracts	NCoR interact with HDAC3	-	Wen et al. 2000
3	*NCoR* expression	Hela nuclear	HDAC3 complex with NCoR and SMRT	-	Li et al. 2000
4	Immunohistochemistry and In situ hybridization	Mouse embryo fibroblasts	NCoR has a critical role in brain development and hematopoietic system	Generation of *NCoR* gene deleted in mice	Jepsen et al. 2000
5	Immunoprecipitations	HeLa cells	NCoR1 has active subunit for HDAC3	-	Guenther et al. 2001
6	Biochemical purification	HeLa nuclear extracts	NCoR mediates DNA methylation	-	Yoon et al. 2003
7	Immunoprecipitation experiments	-	NCoR coactivate BTG1 and influence myoblast differentiation enhances the NCoR1 transcription	-	Busson et al. 2005
8	NMR	DADm mice.	NCoR1 activate HDAC3 in epigenetics regulation of metabolism	Generation of DADm mice	Alenghat et al. 2008
9	Histological analyses	Mice	NCoR1 has an adaptive role in the physiology of muscle	-	Yamamoto et al. 2011
10	Gene expressionanalysis	Members affected with left-sided CHD and normal	*NCoR1* has associated gene in left-sided heart disease	CHD	Hitz et al. 2012
11	Gene expression, chromatin immunoprecipitation	U87 cell lines	Transcriptional activity of NCoR and is involved in regulating pathways for tumour characteristics	Glioblastoma tumour cell	Heldring et al. 2013
12	Histological analysis, Western Blotting, Real- Time PCR.	Mice	*NCoR1* is the principal regulator of thyroid hormone function	-	Shimizu et al. 2014
13	Site-directed mutagenesis	Transgenic Mice	HDAC3 is one part of NCoR1 complexes necessary for second heart field heart development	CHD	Lewandowski et al. 2015
14	Histology	Mice	MeCP2 regulates lipid homeostasis by anchoring the repressor complex NCoR1HDAC3	Rett syndrome	Kyle et al. 2016
15	Cell culture, gene expression array analysis	Healthy prostate and 409 prostate cancer	NCoR1 protein levels were lower in prostate cancer than in healthy prostate	Prostate cancer	Lopez et al. 2016
16	Embryo chromatin immunoprecipitation	Knockout mouse embryos	When NCoR1 is lost, small somite formed	-	Kumar et al. 2016
17	Immunoprecipitation and Western blotting assays	Gastric gastrointestinal stromal tumour	*NCoR1* is tumour suppressors in GIST cells by Smad signalling pathway	Gastric gastroi ntestinala stromal tumour	Wang et al. 2016
18	ChIP-Seq	Mice	NCoR1 is not sufficient to mediate the actions of the unliganded Thyroid hormone receptors (TRs)	Hypothyroidism	Mendoza et al. 2017
19	Chromatinimmunoprecipitation	Cell lines	NCoR1 protein with CHD8 and control negatively in the proliferation of colorectal cancer	Colorectal cancer	Loiselle et al. 2018
20	Affymetrix microarray	Diabetic mice	*NCoR1, ICAM1* and *AK3* are a critical role in heart development	Diabetic	Lin et al. 2018
21	Transcription analysis	autism spectrum disorder	Bifid Uvula due to de novo haploinsufficiency of *NCoR1*	autism spectrum disorder	Sakaguchi et al. 2018
22	Enrichment analyses tool	772 samples from patients	*NCoR1* is a tumour suppressor	Breast and lung tumours	Del et al. 2018

RECOMMENDATIONS

Although the results of this study indicate strong evidence for *NCoR*1 to be an important gene in heart development, this has to be validated in a larger sample size. This will be useful for developing a gene panel for the early detection of VSD and TOF manifestation.

ACKNOWLEDGEMENTS

The researchers are very grateful to the participants, their families and doctors, and technical staffs in various hospitals. The authors also thank Dr. Azim Gholizadeh, Dr. Hadi Esfandi, Dr. Tejas Kalastavadi, Dr. Lingaiah Kusuma, Dr. Chaithra PT and other laboratory colleagues for their support and help during the research work.

NOTES

Table 1: Damaging non-synonymous heterozygous gene variants identified in non-syndromic VSD and TOF in different transcripts of Whole-exome sequence data

*High ranks (0-10%) indicate a gene exhibit more haploinsufficiency

** Probability Loss Intolerance (PLI) from the Exome Aggregation Consortium (ExAC) browser, (>0.9) indicates extremely Loss of Function (LOF), and genes with low PLI scores (PLI $\leq$0.1) are LOF tolerant.

REFERENCES

Bailey P, Downes M, Lau P, Harris J, Chen SL et al. 1999. The nuclear receptor corepressor NCoR regulates differentiation: NCoR directly interacts with MyoD. *Molecular Endocrinology*, 13: 1155-1168.

Bhardwaj R, Rai SK, Yadav AK, Lakhotia S, Agrawal D et al. 2015. Epidemiology of Congenital Heart Disease in India. *Congenital Heart Disease*, 10: 437-446.

Blue GM, Kirk EP, Giannoulatou E, Sholler GF, Dunwoodie SL et al. 2017. Advances in the genetics of congenital heart disease: A clinician's guide. *Journal of the American College of Cardiology*, 69: 859-870.

Cao Y, Wang J, Wei C, Hou Z, Li Y et al. 2016. Genetic variations of NKX2-5 in sporadic atrial septal defect and ventricular septal defect in Chinese Yunnan population. *Gene*, 575: 29-33.

Chelo D, Nguefack F, Menanga AP, Um SN, Gody JC et al. 2016. The Spectrum of heart diseases in children: an echocardiographic study of 1,666 subjects in a pediatric hospital, Yaounde, Cameroon. *Cardiovascular Diagnosis and Therapy*, 6: 10-19.

Chaudhry TA, Younas M, Baig A 2011. Ventricular septal defect and associated complications. *Journal of Pakistan Medical Association*, 61: 1001-1004.

Chen HX, Zhang X, Hou HT, Wang J, Yang Q et al. 2017. Identification of a novel and functional mutation in the TBX5 gene in a patient by screening from 354 patients with an isolated ventricular septal defect. *European Journal of Medical Genetics*, 60: 385-390.

Dinesh SM, Kusuma L, Smitha R, Savitha MR, Krishnamurthy B et al. 2010. Single-nucleotide polymorphisms of NKX2. 5 found in congenital heart disease patients of Mysore, South India. *Genetic Testing and Molecular Biomarkers*, 14: 873-879.

Flaquer A, Baumbach C, Piñero E, Algas FG, de la Fuente Sanchez MA et al. 2013. Genome-wide linkage analysis of congenital heart defects using MOD score analysis identifies two novel loci. *BMC Genetics*, 14: 44-51.

Fu F, Deng Q, Lei TY, Li R, Jing XY et al. 2017. Clinical application of SNP array analysis in fetuses with ventricular septal defects and normal karyotypes. *Archives of Gynaecology and Obstetrics*, 296: 929-940.

Kumar S, Cunningham TJ, Duester G 2016. Nuclear receptor corepressors NCoR1 and NCoR2 (Smrt) are required for retinoic acid-dependent repression of Fgf8 during somitogenesis. *Developmental Biology*, 418: 204-215.

Li C, Sun XN, Chen BY, Zeng MR, Du LJ et al. 2019. Nuclear receptor corepressor 1 represses cardiac hypertrophy. *EMBO Molecular Medicine*, 11: e9127.

Lin N, Cai Y, Zhang L, Chen Y 2018. Identification of key genes associated with congenital heart defects in embryos of diabetic mice. *Molecular medicine reports*, 17: 3697-3707.

Kusuma L, Dinesh SM, Savitha MR, Krishnamurthy B, Narayanappa D et al. 2011. A maiden report on CRELD1 single-nucleotide polymorphism association in congenital heart disease patients of Mysore, South India. *Genetic Testing and Molecular Biomarkers*, 15: 483-487.

Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N 2011. What can exome sequencing do for you? *Journal of Medical Genetics*, 48: 580-589.

Menting ME, Cuypers JA, Opiæ P, Utens EM, Witsenburg M et al. 2015. The unnatural history of the ventricular septal defect: outcome up to 40 years after surgical closure. *Journal of the American College of Cardiology,* 65: 1941-1951.

Meyners C, Mertens M, Wessig P, Meyer Almes FJ 2017. A Fluorescence Lifetime Based Binding Assay for Class IIa Histone Deacetylases. *Chemistry-A European Journal*, 23: 3107-3116.

Muntean I, Togãnel R, Benedek T 2017. Genetics of congenital heart disease: past and present. *Biochemical genetics,* 55: 105-123.

Okubo A, Miyoshi O, Baba K, Takagi M, Tsukamoto K et al. 2004. A novel GATA4 mutation completely segregated with an atrial septal defect in a large Japanese family. *Journal of Medical Genetics*, 41: e97-.e97.

Peyvandi S, Latal B, Miller SP, McQuillen PS 2019. The neonatal brain in critical congenital heart disease:

insights and future directions. *Neuroimage*, 185: 776-782.

Phelps MP, Bailey JN, Vleeshouwer-Neumann T, Chen EY 2016. CRISPR screen identifies the NCoR/HDAC3 complex as a major suppressor of differentiation in rhabdomyosarcoma. *Proceedings of the National Academy of Sciences*, 113: 15090-15095.

Rosenfeld MG, Lunyak VV, Glass CK 2006. Sensors and signals: A coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes and Development*, 20: 1405-1428.

Saxena A, Mehta A, Sharma M, Salhan S, Kalaivani M et al. 2016. Birth prevalence of congenital heart disease: A cross-sectional observational study from North India. *Annals of Pediatric Cardiology*, 9: 205-209.

Smitha R, Karat SC, Narayanappa D, Krishnamurthy B, Prasanth SN et al. 2006. Prevalence of congenital heart diseases in Mysore. *Indian Journal of Human Genetics*, 12: 11-16.

Tao H, Song ZY, Ding XS, Yang JJ, Shi KH, Li J 2018. Epigenetic signatures in cardiac fibrosis, special emphasis on DNA methylation and histone modification. *Heart Failure Reviews*, 23: 789-799.

Wen YD, Perissi V, Staszewski LM, Yang WM, Krones A et al. 2000. The histone deacetylase-3 complex contains nuclear receptor corepressors. *Proceedings of the National Academy of Sciences*, 97: 7202-7207.

Paper received for publication in August, 2019
Paper accepted for publication in January, 2020