

RESEARCH

Open Access



Identification and functional analysis of variants of *MYH6* gene promoter in isolated ventricular septal defects

Ji-Yang Zuo^{1,2}, Huan-Xin Chen¹, Zhi-Gang Liu¹, Qin Yang¹ and Guo-Wei He^{1*} 

Abstract

Background: Ventricular septal defect is the most common form of congenital heart diseases. *MYH6* gene has a critical effect on the growth and development of the heart but the variants in the promoter of *MYH6* is unknown.

Patients and methods: In 604 of the subjects (311 isolated and sporadic ventricular septal defect patients and 293 healthy controls), DNA was extracted from blood samples and *MYH6* gene promoter region variants were analyzed by sequencing. Further functional verification was performed by cellular experiments using dual luciferase reporter gene analysis, electrophoretic mobility shift assays, and bioinformatics analysis.

Results: Nine variants were identified in the *MYH6* gene promoter and two of those variants [g.4085G>C(rs1222539675) and g.4716G>A(rs377648095)] were only found in the ventricular septal defect patients. Cellular function experiments showed that these two variants reduced the transcriptional activity of the *MYH6* gene promoter ($p < 0.001$). Further analysis with online JASPAR database suggests that these variants may alter a set of putative transcription factor binding sites that possibly lead to changes in myosin subunit expression and ventricular septal defect formation.

Conclusions: Our study for the first time identifies variants in the promoter region of the *MYH6* gene in Chinese patients with isolated and sporadic ventricular septal defect. These variants significantly reduced *MYH6* gene expression and affected transcription factor binding sites and therefore are pathogenic. The present study provides new insights in the role of the *MYH6* gene promoter region to better understand the genetic basis of VSD formation.

Keywords: Ventricular septal defect, *MYH6*, Genetic, Variants, Congenital heart disease

Introduction

Congenital heart defects (CHDs), characterized by structural and conduction abnormalities, are congenital anomalies that pose a serious threat to the health of

children's lives and represent the most prevalent congenital birth defects, accounting for approximately one-third of all congenital defects [1, 2]. CHDs affects 9–18 out of every 1000 live births [3], and nearly 300,000 people die of congenital heart disease each year, 70% of whom are children under the age of one [4]. Advances in the treatment of CHDs have improved the survival rate of CHD patients and have largely led to an increase in its prevalence in older children and adults, who remain at high risk for continued disease burden after surgical repair, with the possibility of atrial arrhythmias, pulmonary hypertension, and other complications requiring repeat

Guo-Wei He: Distinguished Professor of Tianjin University, China. Foreign Correspondence Member: The National Academy of Medicine, France.

*Correspondence: gwhezj@163.com; gwhe@tju.edu.cn

¹ Department of Cardiovascular Surgery, The Institute of Cardiovascular Diseases, TEDA International Cardiovascular Hospital, Tianjin University & Chinese Academy of Medical Sciences, No. 61, the 3rd Ave., Tianjin 300457, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

surgery, subsequently increasing the health expenditure and medical burden on society [5]. Therefore, it has become increasingly important to unravel the pathogenesis of CHDs. Previous studies have shown that genetic, epigenetic and environmental factors are causative factors of CHDs [3, 6, 7].

Ventricular septal defect (VSD) is the most common CHD, accounting for 30–40% of all CHDs [8, 9]. The incidence of this defect varies with age due to the varying sensitivity of screening techniques and the fact that many small malformations present at birth disappear shortly thereafter. The incidence of isolated VSDs by high-sensitivity color Doppler echocardiography screening has been reported to be 5.7% in preterm infants and 1.1–5.3% in full-term infants [10]. However, despite efforts to unravel the mechanisms of VSD formation, the exact mechanisms remain largely unknown.

MYH6 (OMIM: 160710) is located on chromosome 14q11.2, and the protein it expresses is one of two myosin heavy chain isoforms expressed by the myocardium that mark cardiac progenitor cells and are involved in a variety of pluripotent cardiac cell lineages during embryonic development, being an important protein component in the formation of cardiac structure [11, 12]. Its potential impact on many forms of CHDs and cardiomyopathies, such as hypertrophic cardiomyopathy, dilated cardiomyopathy, sick sinus node syndrome, atrial septal defect, pulmonary atresia, and VSD has been reported [13–16]. Single nucleotide polymorphisms in *MYH6* have been found to be potentially associated with CHD [17, 18], whereas no reports have focused on variants in the promoter region of the *MYH6* gene causing CHD. Recent studies on gene promoter regions have shown that mutations and variations in promoter regions can alter the expression of genes, which may lead to related diseases [19–21]. In view of the fact that MYH6 protein is an important subunit of myosin during cardiac development, we hypothesized that variants in the *MYH6* promoter may change the expression of *MYH6* and may be involved in the formation of VSD. To test this hypothesis, we designed the present study to genetically and functionally analyze the DNA sequences of the *MYH6* promoter region in Chinese patients with isolated and sporadic VSD compared to healthy controls.

Patients, materials and methods

Study participants included

This study enrolled 604 subjects. A total of 464 Chinese patients with isolated and sporadic VSD underwent repair surgery from January 2018 to August 2019 at the Department of Cardiovascular Surgery, TEDA International Cardiovascular Hospital, Tianjin, China. A total of 311 VSD patients (167 males and 144 females; age

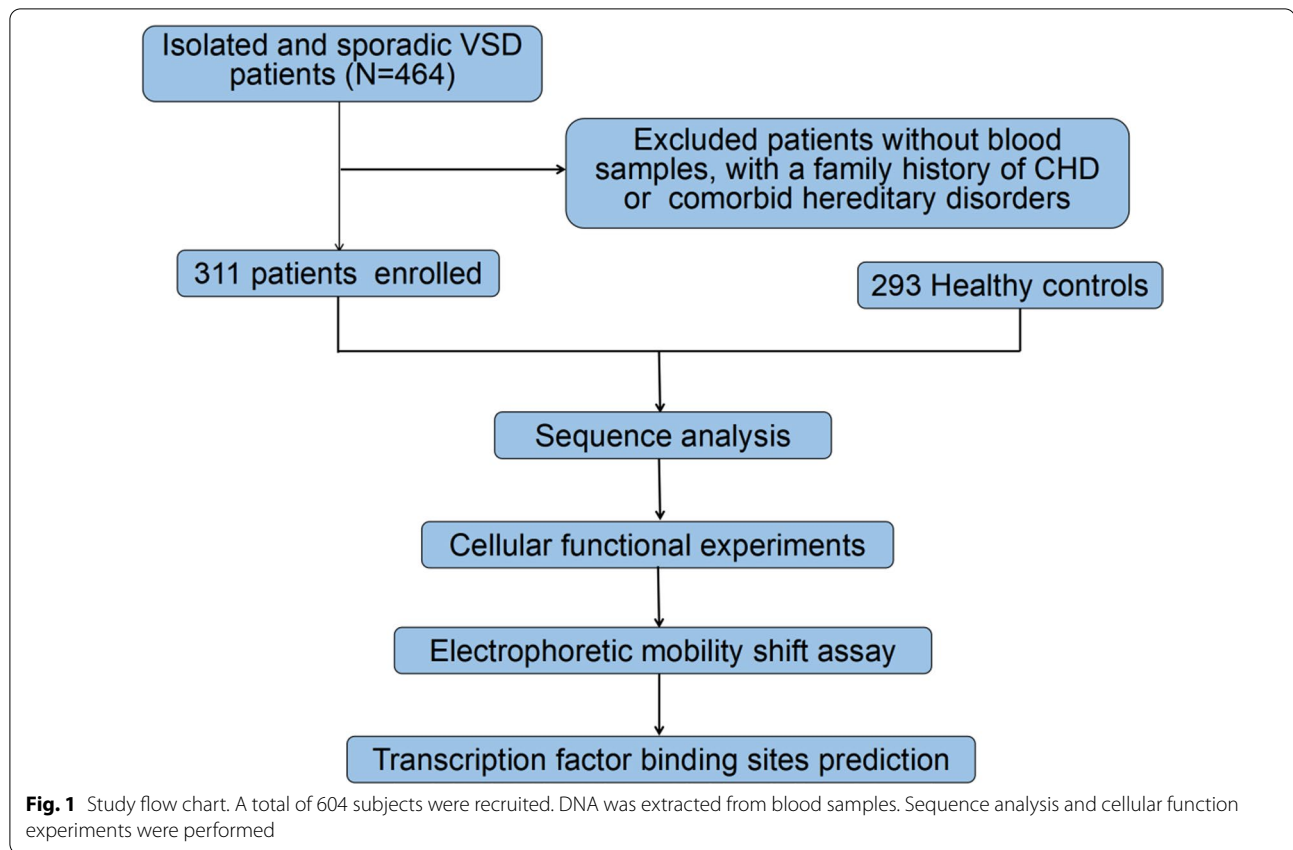
range: 1 month to 12 years) were enrolled in this study. The exclusion criteria were patients without blood samples, or with a family history of CHD, or with other comorbid hereditary disorders. The blood samples from those patients were collected. In addition, blood samples from 293 healthy controls (166 males and 127 females; age range: 1 month to 12 years) were also collected. The healthy controls were from heart disease screening program and body check program. All healthy subjects were confirmed free of CHD or other diseases and family history of genetic disorders by clinical screening including echocardiography. This work was performed in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of TEDA International Cardiovascular Hospital. Written consent forms were informed and signed by the parents or guardians of the participants. The flow chart of the study is shown in Fig. 1.

Sequence analysis

Fasting venous blood was collected and genomic DNA was extracted from peripheral leukocytes using the RelaxGene Blood DNA System kit (TianGen, Beijing, China). The sequence of the *MYH6* gene promoter was obtained from the GenBank database (NCBI, NG_023444.1) and was amplified from human genomic DNA. The design of the polymerase chain reaction (PCR) primers was based on the sequence of the human *MYH6* gene. The gene promoter is located upstream of the coding region starting at –1500 bp to –2000 bp, and the specific promoter sequence of *MYH6* was selected by BLAST in GenBank database (1638 bp, from –1791 bp to –154 bp to the transcription start site). Amplification conditions were: 95 °C for 5 min, followed by 35 cycles (denaturation at 95 °C for 30 s, annealing at 60 °C for 30 s, extension at 72 °C for 105 s), and a final extension at 72 °C for 10 min. The PCR products were sequenced in both directions for alignment and comparison with the wild-type *MYH6* gene promoter sequence. Primers for Sanger sequencing are shown in Table 1. DNA sequences were then compared to the wild-type *MYH6* gene promoter and regulatory variants, including single nucleotide polymorphisms (SNPs), were identified.

Plasmid constructs, cell culture, and transfection

To investigate the effect of variants on promoter activity, the wild-type and variant *MYH6* gene promoter fragments generated by PCR were inserted into the KpnI and SacI sites of pGL6-basic to construct gene expression vectors, which were verified by Sanger sequence analysis. The PCR primers with KpnI and SacI sites are listed in Table 1. The plasmids were extracted using Plasmid Mini Preparation Kit (Beyotime Biotechnology, Shanghai,

**Table 1** List of primers used in this study

Primers name	Sequences	Location
PCR primers		
MYH6-F ₁	5'-GGGGCCTCGAGTAACCTAGA-3'	3209–3228
MYH6-R ₁	5'-CCCCCTGATTGCCCAAGAA-3'	4846–4827
Sequencing primers		
MYH6-F ₁	5'-GGGGCCTCGAGTAACCTAGA-3'	3209–3228
MYH6-R ₂	5'-CCCCCTGATTGCCCAAGAA-3'	4846–4827
MYH6-R ₁	5'-GACTTGACCGTGTCTGTGC-3'	4135–4154
Primers containing restriction sites		
MYH6-KpnI ^a	5'- CGG GGTACCGGGCCTCGAGTAACCTAGA-3'	
MYH6-SacI ^a	5'- AAA AGTACTCCCCCTGATTGCCCAAGAA-3'	
The double-stranded biotinylated oligonucleotides for the EMSA		
g.4085G>C-F	5'-CACACTTACCCACTCAAGCTTG(C/G)AGATTCTTTTCTACTTTC-3'	
g.4085G>C-R	5'-GAAAGTAGAAAAGAAATCT(G/C)CAAGCTTGAGTGGGTAAGTGTG-3'	
g.4716G>A-F	5'-AAGACAGGACCTTCTCA(C/T)ACCGCCTCTCCCACCCT-3'	
g.4716G>A-R	5'-AGGGTGGGAGAGGCGGT(G/A)TGAGAAGTCCTGTCTT-3'	
Nucleotide sequences for JASPAR prediction of TFBS		
g.4085G>C-F	5'-CACACTTACCCACTCAAGCTTG(C/G)AGATTCTTTTCTACTTTC-3'	
g.4716G>A-F	5'-AAGACAGGACCTTCTCA(C/T)ACCGCCTCTCCCACCCT-3'	

PCR primers are designed based on the genomic DNA sequence of the *MYH6* gene (NG_023444.1). The transcription start site is at the position of 5,001 (+1)

EMSA, electrophoretic mobility shift assay; TFBS, transcription factor binding sites; F, forward; R, reverse

^a Protective bases are presented in bold

China) according to the instructions. These expression vector plasmids were then transfected separately with pRL-SV40 (sea kidney luciferase reporter plasmid) in a 5:1 ratio into HEK-293 and HL-1 cells, which were cultured in Minimum Essential Medium supplemented with 10% fetal bovine serum. For cell transfection analysis, 1×10^6 cells were transfected with 2.5 ng of plasmid DNA, and all groups were transfected together in six-well plates and subsequently examined for dual luciferase activity.

Analysis with the dual-luciferase reporter assay

After 48 h of transfection, the cells were collected and lysed sufficiently to retain the lysate. Dual luciferase activity was measured using a dual luciferase reporter gene assay system (Beyotime Biotechnology, Shanghai, China). The empty vector pGL6-basic and blank were used as negative controls and pRL-SV40 was used as an internal control to correct for transfection efficiency. The results are presented in terms of the relative fold change of these construction vectors compared to the wild-type *MYH6* gene promoter in the expression vector. All experiments were repeated five times independently.

Nuclear extracts preparation and electrophoretic mobility shift assay (EMSA)

To examine the effect of *MYH6* gene regulatory variants on transcription factor binding sites (TFBS), nuclear protein extracts of HEK-293 and HL-1 were prepared using the Nuclear and Cytoplasmic Protein Extraction Kit (Beyotime Biotechnology, Shanghai, China) according to the manufacturer's instructions to EMSA. Protein concentrations were quantified using the Enhanced BCA Protein Assay Kit (Beyotime Biotechnology, Shanghai, China). Biotinylated double-stranded oligonucleotides containing wild-type or variant DNA sequences in the *MYH6* gene promoter were used as probes. EMSA was performed using the chemiluminescent EMSA kit (Beyotime Biotechnology, Shanghai, China) using equal amounts of probes (0.25 pM) and nuclear extracts (2.0 µg) according to the manufacturer's protocol. The EMSA oligonucleotide sequences are shown in Table 1.

TFBS and interacting proteins prediction

JASPAR (<https://jaspar.genereg.net/>) [22] (version 9) was used to predict the possible binding transcription factors (TFs) [23]. To further investigate whether *MYH6* gene promoter variants disrupt or generate TFBS, we used JASPAR to generate a list of all potentially affected binding sites through the identified variants with a relative profile score threshold set at 85%. The wild-type and variant nucleotide sequences are shown in Table 1. STRING (<https://cn.string-db.org/>) is an online

searchable database of known protein–protein interaction (PPI) [24]. By using the protein names provided by the user, the interacting proteins can be more accurately predicted. We used STING to find proteins that may be affected by mutations in the *MYH6* gene with setting the organism to "Homo sapiens" and setting the minimum required interaction score ≥ 0.700 . We then imported the resulting PPI maps into Cytoscape 3.9.1 software for network topology analysis [25].

Statistical analysis

The quantitative data were expressed as mean \pm SD and compared by one-way analysis of variance. The frequency of DNA sequence variants was compared with that of VSD patients and healthy controls using the Pearson's chi-square. All statistical analyses were performed with SPSS 25.0 software. $p < 0.05$ was considered statistically significant.

Results

Variants found in VSD patients by DNA sequence

In the 604 Chinese subjects recruited for this study, nine variants were identified by Sanger sequencing. All of these variants are summarized in the Table 2. The locations are shown in Fig. 2A and their sequencing chromatograms are shown in Fig. 2B. Among these 9 variants, 2 single nucleotide variants [g.4085G>C(rs1222539675) and g.4716G>A(rs377648095)] were found only in VSD patients but the other 7 variants were found in both VSD patients and healthy controls (Table 2). Importantly, in the NCBI SNP database, the above two variants only found in the VSD patients had allele frequencies < 0.0001 [g.4085G>C(rs1222539675) variant rate C = 0.000019 and g.4716G>A(rs377648095) variant rate T = 0.000011], suggesting the possible pathogenic role of these variants. Therefore, these two variants were further studied in cellular function experiments to reveal the pathological role. In contrast, the other seven variants found in both VSD patients and controls were excluded from further studies.

Regulatory variants alter the activity of the *MYH6* gene promoter

In order to examine the effect of the variants in the *MYH6* gene promoter on promoter activity, the wild-type and variant *MYH6* gene promoters were cloned into the firefly luciferase reporter plasmid (pGL6-basic) reporter vector to generate constructed expression vectors, including blank, empty pGL6-basic (negative control), pGL6-WT (wild-type *MYH6* gene promoter), pGL6-4085C (g.4085G>C) and pGL6-4716A (g.4716G>A), respectively. These constructed vectors were then transfected into HEK-293 and HL-1 cells and dual luciferase activity was measured. The

Table 2 Variants within the *MYH6* gene promoter in patients with VSD

Variations	Position ^a	Genotypes	VSD ^b	Controls ^b	Frequency ^c	P value
g.4085G>C(rs1222539675)	– 915	G>C	1	0	C = 0.000019	–
g.4716G>A(rs377648095)	– 284	G>A	1	0	A = 0.000011	–
g.3285A>G(rs17091776)	– 1715	A>G	56	42	G = 0.1048	0.221
g.3557G>T(rs191392051)	– 1443	G>T	18	10	T = 0.0038	0.165
g.3726A>G(rs178648)	– 1244	A>G	7	2	G = 0.0046	0.178
g.3816C>A(rs138953808)	– 1184	C>A	8	9	A = 0.0224	0.711
g.3931del(rs377182175)	– 1069	delC	5	4	delC = 0.0014	1.0
g.4208A>G(rs9788443)	– 972	A>G	63	45	G = 0.1767	0.116
g.4387T>C(rs73587609)	– 613	T>C	62	46	C = 0.0477	0.175

^a Variants are located upstream (–) to the transcription start site at the position of 5,001 (+) of the *MYH6* gene (NG_023444.1)

^b Allele frequency in groups. VSD, ventricular septal defect

^c The allele frequency was obtained from NCBI dbSNP database

transcriptional activity of the blank and negative control groups was much lower than the other three groups, nearly 0. The two variants, g.4085G>C (rs1222539675) and g.4716G>A (rs377648095) found in VSD patients, significantly reduced the gene promoter expression activity and decreased luciferase expression compared to the wild-type *MYH6* gene promoter ($p < 0.001$) in HEK-293 (Fig. 3A) and HL-1 cells (Fig. 3B). The color Doppler echocardiography of the patients with these two variants is shown in Fig. 3C.

EMSA analysis

To experimentally investigate whether regulatory variants affect TF binding, EMSA was tested using wild-type or variant biotin-labeled probes. Regulatory variants found only in VSD patients were examined. The biotinylated oligonucleotides of EMSA are shown in Table 1. As shown in Fig. 4A, the variant g.4085G>C enhances the binding of TFs. In contrast, variant g.4716G>A significantly abolished the binding of TFs and created binding sites for new TFs in HEK-293 cells. Similarly, in HL-1 cells (Fig. 4B), the variant g.4085G>C and variant g.4716G>A enhance the binding of TFs in HL-1 cells (The original figures are provided in Additional file 1: Fig. S1). The brightness of the band may reflect the binding ability of TFs [26].

The TFBS affected by regulatory variants and the PPI of MYH6

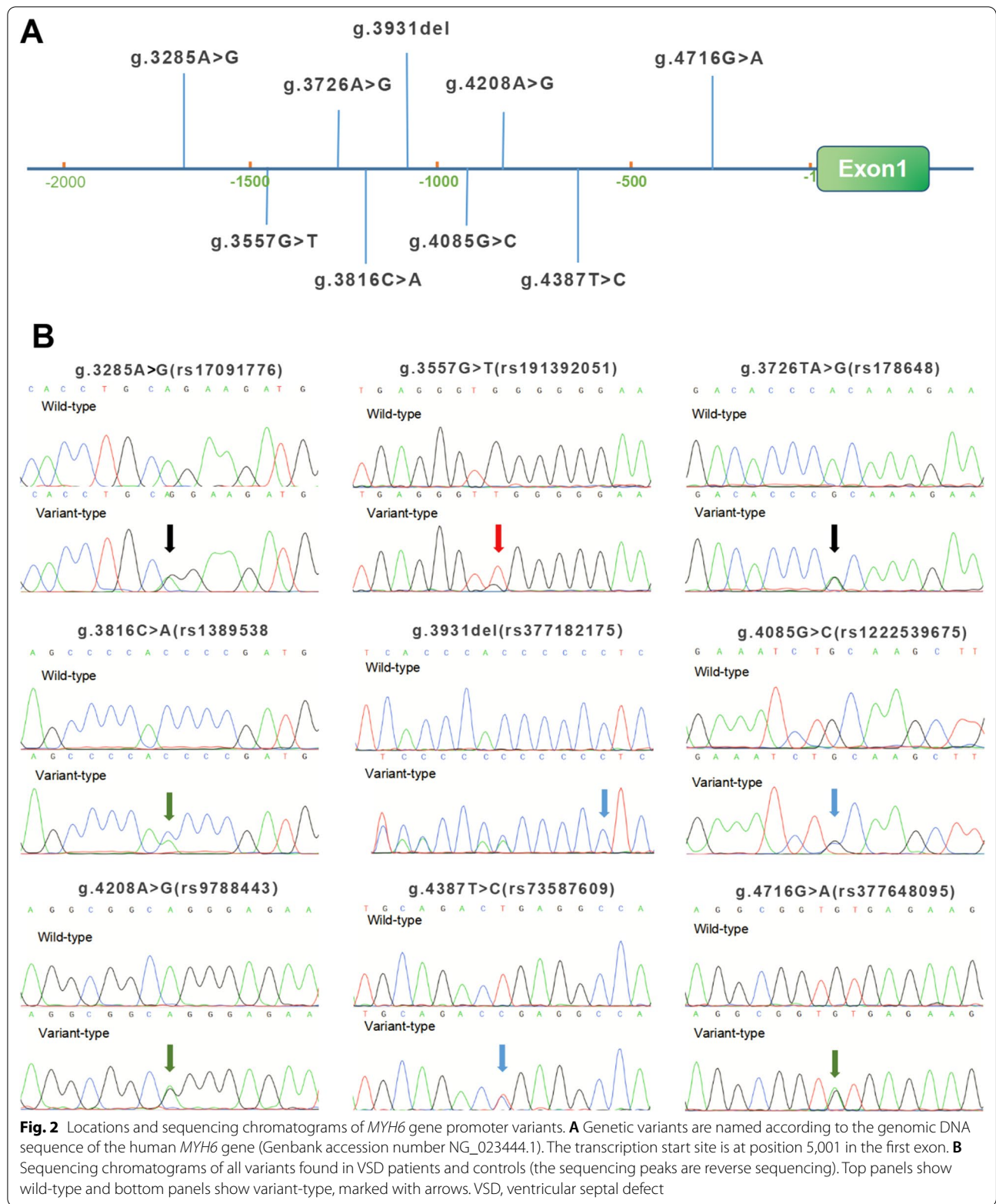
Based on the JASPAR core TF database, we predicted potential TFBS that may be disrupted or generated by variants in the *MYH6* gene promoter. The results showed that various TFs play important roles in cardiogenesis, such as *E2F1*, *GATA4*, *NKX2-5*, *GATA2*, *GATA1*, *GATA6*, etc. As shown in the Table 3, variant g.4085G>C may generate binding sites for *NFIX*, *PAX6*, *Gfi1* and *NFIC*

whereas variant g.4716G>A may produce binding sites for *TEAD1* and *TEAD3* and disrupt binding sites for *SP2*, *TBX4*, *TBX1*, *TBX15*, *SOX18*, *TBX6*, and *TBX18*. Based on these analyses by using the JASPAR database, cellular function experiments, and previous studies, we established a schema to describe the possible effects of variants in the promoter region of the *MYH6* gene on the development of VSD as shown in Fig. 5. The construction of PPI networks using STRING was analyzed by Cytoscape 3.9.1, and a total of 31 nodes and 300 edges were obtained (Fig. 4C). The nodes of the network graph represent the protein names and the edges represent their interactions. The top ten proteins that interacted most strongly with MYH6 protein were TNNI1, MYBPC3, MYL2, ACTC1, ACTN2, MYL1, TNNT2, TPM1, and TNNI3.

Discussion

The present study focused on variants in the *MYH6* gene promoter in the Han Chinese patients with isolated and sporadic VSD and has for the first time found that (1) among the 9 regulatory variants identified within the promoter region of the *MYH6* gene in the 604 human subjects, two variants [g.4085G>C(rs1222539675) and g.4716G>A(rs377648095)] were only detected in VSD patients in the study population; (2) these two variants cause cellular functional changes demonstrated in dual luciferase activity and EMSA; (3) These variants may cause alterations in TFBS and also affect proteins that interact with the MYH6 protein, resulting in altered function and ultimately leading to VSD.

MYH6 gene encodes the α heavy chain subunit of cardiac myosin [27], which is expressed in the heart from fetal life to adulthood and is actively involved in muscle force production and contraction [11, 12]. Altered transcriptional activity of the *MYH6* gene has been reported to affect the formation of myocardial protofibrils in



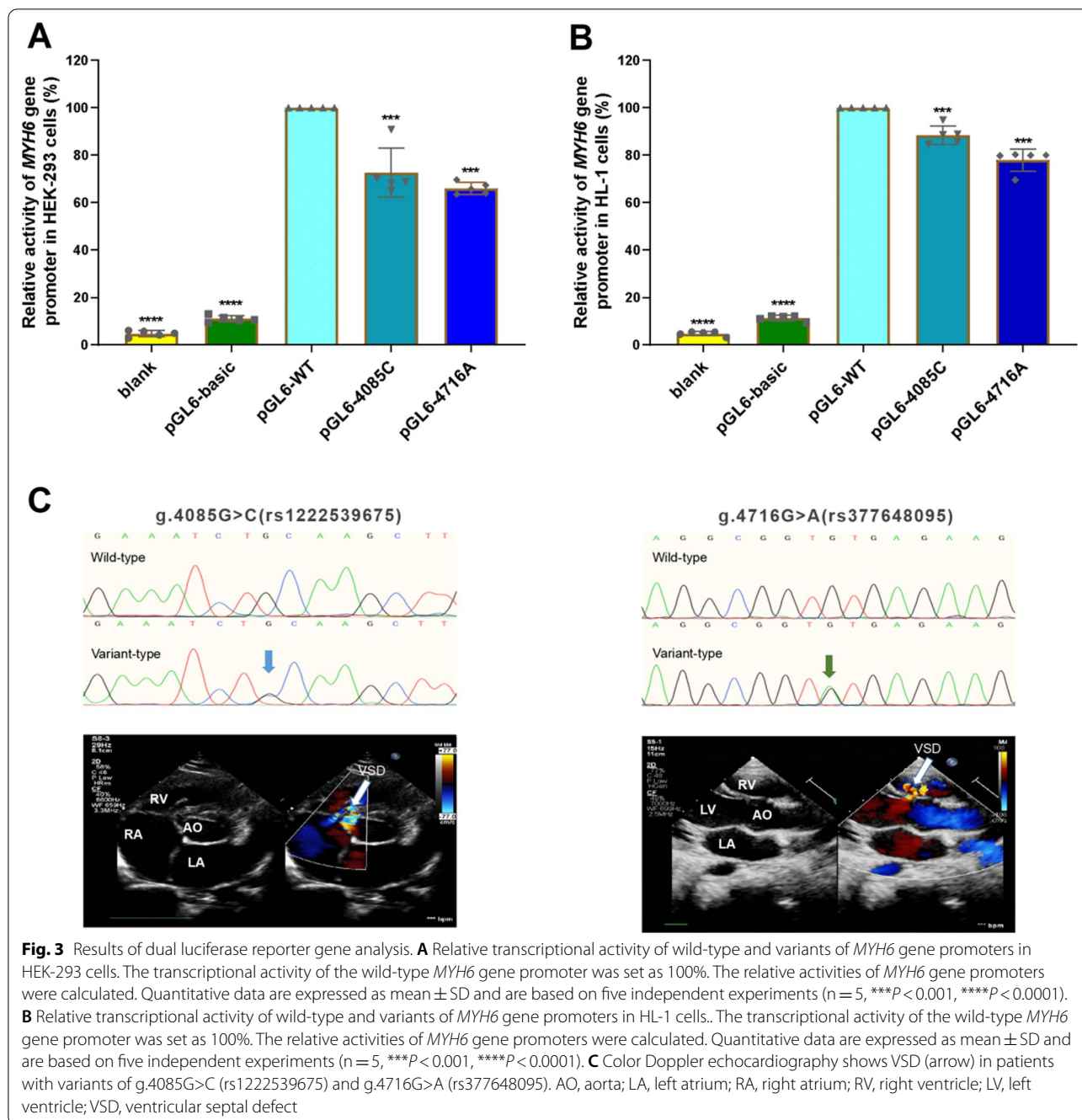


Fig. 3 Results of dual luciferase reporter gene analysis. **A** Relative transcriptional activity of wild-type and variants of *MYH6* gene promoters in HEK-293 cells. The transcriptional activity of the wild-type *MYH6* gene promoter was set as 100%. The relative activities of *MYH6* gene promoters were calculated. Quantitative data are expressed as mean \pm SD and are based on five independent experiments ($n = 5$, $***P < 0.001$, $****P < 0.0001$). **B** Relative transcriptional activity of wild-type and variants of *MYH6* gene promoters in HL-1 cells. The transcriptional activity of the wild-type *MYH6* gene promoter was set as 100%. The relative activities of *MYH6* gene promoters were calculated. Quantitative data are expressed as mean \pm SD and are based on five independent experiments ($n = 5$, $***P < 0.001$, $****P < 0.0001$). **C** Color Doppler echocardiography shows VSD (arrow) in patients with variants of g.4085G>C (rs1222539675) and g.4716G>A (rs377648095). AO, aorta; LA, left atrium; RA, right atrium; RV, right ventricle; LV, left ventricle; VSD, ventricular septal defect

animal experiments, resulting in the inability to form heart valves and the formation of heart trabeculae or even death [28, 29]. Recent studies have shown that genetic polymorphisms in *MYH6* are associated with risk of VSD [17, 18, 30]. Further, there is growing evidence that variants in the gene promoter play a key role in cardiac transcriptional regulation [19–21, 31, 32]. Based on the STRING database and a previous study [33], we found that the major proteins interacting with *MYH6*

protein are basically involved in myocardial development. Alterations in *MYH6* protein expression may affect the interaction with other proteins that may have important implications for myocardial development. Therefore, the low-expression of *MYH6* protein may lead to defects in myocardial formation.

It has been reported that variants in the *MYH6* gene may affect the formation of the myosin complex [34, 35]. However, mutations or variants in the promoter

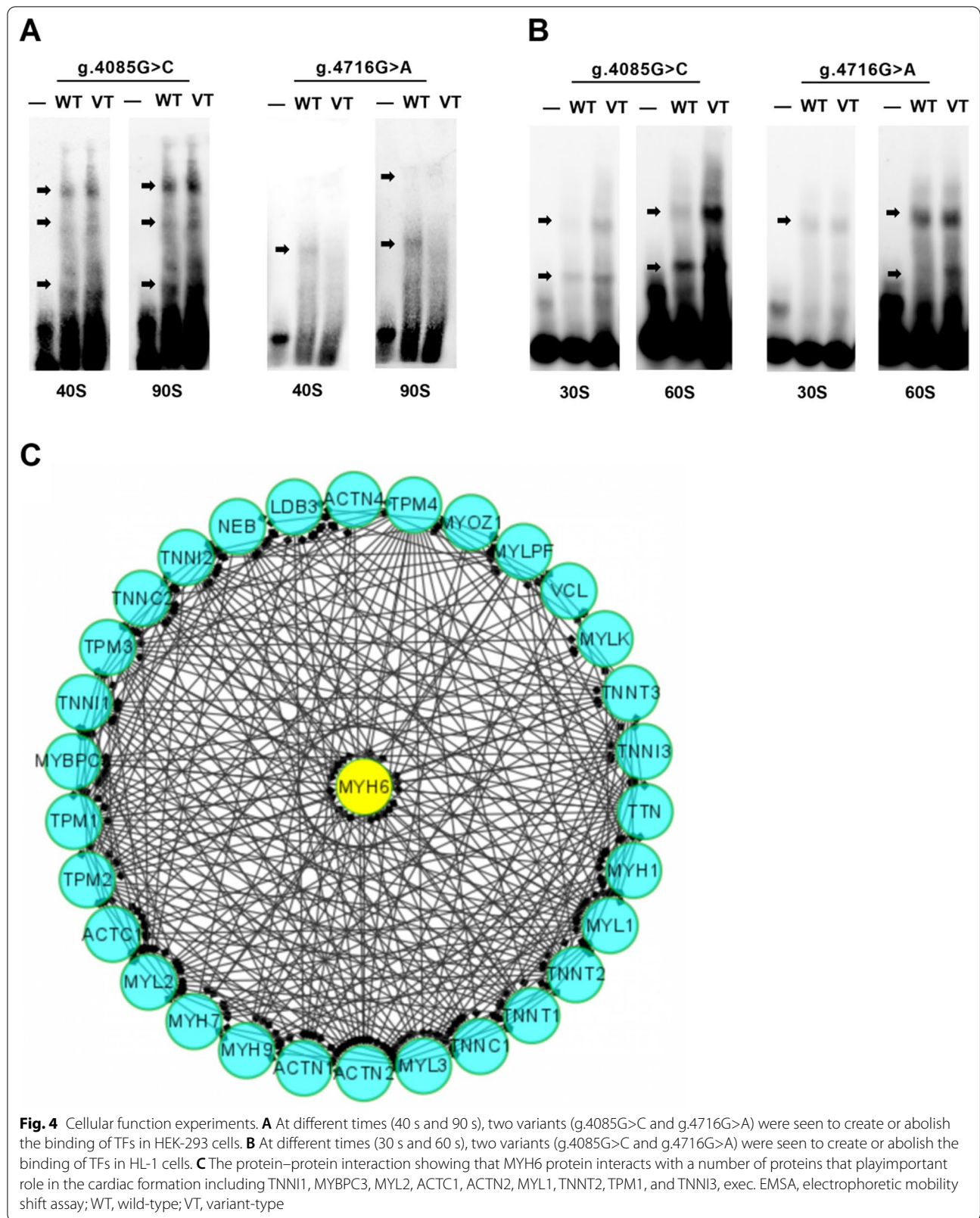


Fig. 4 Cellular function experiments. **A** At different times (40 s and 90 s), two variants (g.4085G>C and g.4716G>A) were seen to create or abolish the binding of TFs in HEK-293 cells. **B** At different times (30 s and 60 s), two variants (g.4085G>C and g.4716G>A) were seen to create or abolish the binding of TFs in HL-1 cells. **C** The protein–protein interaction showing that MYH6 protein interacts with a number of proteins that play important role in the cardiac formation including TNNI1, MYBPC3, MYL2, ACTC1, ACTN2, MYL1, TNNT2, TPM1, and TNNI3, exec. EMSA, electrophoretic mobility shift assay; WT, wild-type; VT, variant-type

Table 3 TFBS affected by variants (predicted by the JASPAR database)

	Wild ^a	Variant ^b	Disrupt	Create
g.4085G>C	NKX2-5, NKX3-2, ZNF354C, FOXC1, NKX3-1, GATA3, GATA2, GATA1, NR2E3, NFAT5, EN2, NKX2-8, RHOXF1, HIC2, ISL2, BARHL1, MSANTD3, HOXB6, DLX1, DIX4, GATA6, DLX6, DIX3, BARX1, FOXP3, DIX2, NKX2-2, NKX2-4, SOX18, ZNF784, ZIM3	NKX2-5, NKX3-2, ZNF354C, FOXC1, NKX3-1, GATA3, GATA2, GF11, GATA1, PAX6, NR2E3, NFAT5, NFIC, EN2, NKX2-8, RHOXF1, HIC2, NFIX, ISL2, BARHL1, MSANTD3, HOXB6, DLX1, DIX4, GATA6, DLX6, DIX3, BARX1, FOXP3, DIX2, NKX2-2, NKX2-4, SOX18, ZNF784, ZIM3	–	NFIX, PAX6, Gf11, NFIC
g.4716G>A	E2F1, KLF4, SP1, SPI1, KLF5, MEIS1, E2F6, EGR1, GATA4, KLF1, E2F4, SP2, MGA, MEIS3, TBX4, TBX5, TBX1, TBX15, SP8, SP3, KLF14, NR2C2, NR4A1, ZNF75D, SOX18, WT1, TBX6, TBX18, ZNF281, PRDM9, SP9, TBX3	E2F1, KLF4, SP1, SPI1, KLF5, MEIS1, E2F6, EGR1, GATA4, KLF1, E2F4, MGA, MEIS3, TBX5, SP8, SP3, KLF14, NR2C2, NR4A1, ZNF75D, WT1, ZNF281, PRDM9, SP9, TBX3, TEAD1, TEAD3	SP2, SOX18, TBX1, TBX4, TBX6, TBX15, TBX18	TEAD1, TEAD3

TFBS, transcription factor binding sites

^a Wild-type possessed TFBS

^b Variant-type possessed TFBS

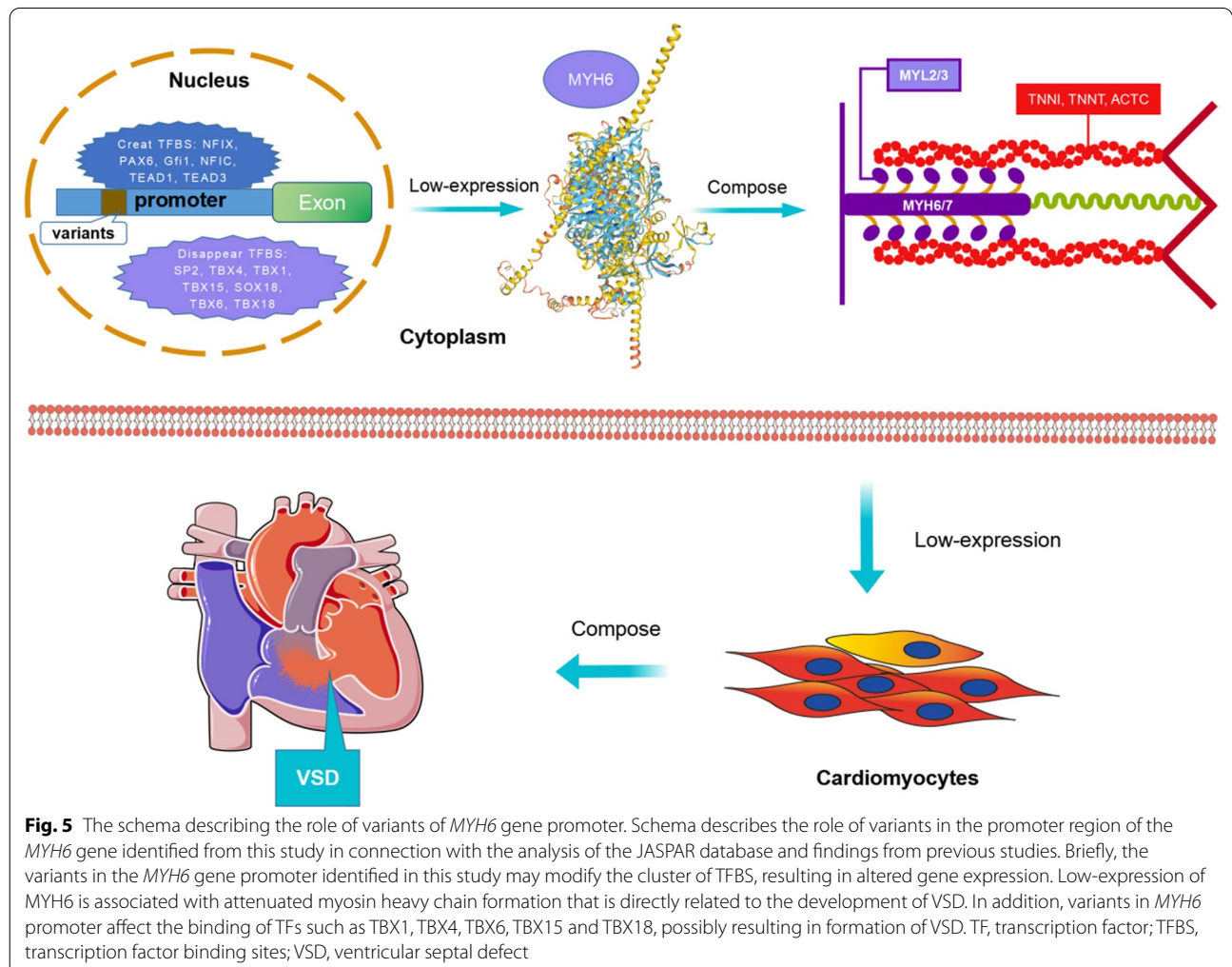


Fig. 5 The schema describing the role of variants of *MYH6* gene promoter. Schema describes the role of variants in the promoter region of the *MYH6* gene identified from this study in connection with the analysis of the JASPAR database and findings from previous studies. Briefly, the variants in the *MYH6* gene promoter identified in this study may modify the cluster of TFBS, resulting in altered gene expression. Low-expression of *MYH6* is associated with attenuated myosin heavy chain formation that is directly related to the development of VSD. In addition, variants in *MYH6* promoter affect the binding of TFs such as TBX1, TBX4, TBX6, TBX15 and TBX18, possibly resulting in formation of VSD. TF, transcription factor; TFBS, transcription factor binding sites; VSD, ventricular septal defect

region of the *MYH6* gene have not been reported. In our study population (n=604), *MYH6* gene regulatory variants were identified in VSD patients with a frequency of 0.64% (2/311) and the frequency in the total study population was 0.33% (2/604). Further, in this study we demonstrated that these two variants in the promoter region of the *MYH6* gene significantly lowered the transcriptional activity of *MYH6*, demonstrated by the changes induced by these variants in dual fluorescein reporter analysis. As well known, changes of the activity of *MYH6* affect the formation of myosin heavy chains [27].

In the present study, EMSA experiments showed that two variants [g.4085G>C (rs1222539675) and g.4716G>A (rs377648095)] only found in VSD patients had altered binding affinity to TFs. As reported, changes of binding affinity to TFs are critical in the transcriptional regulation of gene expression that leads to disease [19–21, 36]. Therefore, our study suggests that variants in the *MYH6* gene promoter may alter the TFBS cluster bound in the promoter, thereby reducing *MYH6* transcript levels that may be involved in the development of VSD through specific pathways.

Based on the JASPAR database analysis, two variants [g.4085G>C (rs1222539675) and g.4716G>A (rs377648095)] have been predicted to have TFBS generation or disappearance. Among the TFBS derived from the analysis, it has been demonstrated that *PAX6* inhibited the differentiation of cardiac fibroblasts [37]. In contrast, most of the TFBS that existed in the wild type but were deleted in the presence of the variants play essential roles in cardiovascular development, such as *SP2* and *SOX18* [38, 39]. T-box family TFs play essential roles in the development of the embryonic heart [40, 41]. It was predicted that variant g.4716G>A may attenuate the bindings of *TBX1*, *TBX4*, *TBX6*, *TBX15* and *TBX18* to the promoter region of the *MYH6* gene. In fact, these genes are differentially expressed at various stages of heart development, participating in the formation of the heart [42–46]. In summary, the variants in the promoter of *MYH6* gene identified in this study may modify the TFBS clusters in the promoter and interfere with the formation of transcriptional regulatory complexes with related TFs, leading to the formation of VSD. Figure 5 is a schema that illustrates these results.

As to the variants other than the two variants that were only found in the VSD cases (Table 2), although some of them have low frequency in the NCBI dbSNP database (<1%), whether they may also affect the expression needs to be further studied. In this study, due to the fact that they were also found in the controls, we excluded them from further cellular experiments.

Limitation

Although this study investigated samples from VSD patients and control subjects and the variants found had further functional tests at cellular level to reveal the potential pathogenic role, further animal experiments are needed to confirm the role of these variants in the development of VSD. This will be considered in our future studies.

Conclusions

Our study identified nine variants in the promoter region of the *MYH6* gene in Chinese patients with isolated and sporadic VSD, two of which were found only in VSD patients. Furthermore, cellular function experiments, EMSA and bioinformatics analyses showed that these variants remarkably altered the expression of the *MYH6* gene and affected the binding of TFs that may contribute to the development of VSD. Therefore, the present study may provide new insights into the role of gene promoter regions to better understand the genetic basis of VSD formulation and facilitate further studies on the mechanisms of VSD formation.

Abbreviations

CHD: Congenital heart defect; VSD: Ventricular septal defect; PCR: Polymerase chain reaction; SNP: Single nucleotide polymorphism; TFBS: Transcription factor binding sites; TF: Transcription factor; EMSA: Electrophoretic mobility shift assay; PPI: Protein–protein interaction; WT: Wild-type; VT: Variant-type; AO: Aorta; LA: Left atrium; RA: Right atrium; RV: Right ventricle; LV: Left ventricle.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12920-022-01365-y>.

Additional file 1. Fig. S1: The original images of the EMSA of the two variants (g.4085G>C and g.4716G>A) with nuclear proteins from HEK-293 cells and HL-1 cells. EMSA, electrophoretic mobility shift assay; WT, wild-type; VT, variant-type.

Acknowledgements

We thank the patients and their family members for cooperation. The assistance of Nursing staff at the Division of Pediatric Cardiac Surgery, Department of Cardiovascular Surgery is gratefully acknowledged.

Author contributions

GWH conceptualized the project. GWH and HXC designed the study. GWH obtained the funding for this project. JYZ, HXC performed experiments. JYZ, HXC, and GWH analyzed experimental data. JYZ, HXC, ZGL, and GWH collected the human blood samples. QY discussed the findings and participated in the project. JYZ, and GWH wrote the manuscript. GWH supervised the whole project. All authors read and approved the final manuscript.

Funding

This work was funded by the National Natural Science Foundation of China [82170353 & 81870288]; the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences [2020-PT310-007]; Tianjin Municipal and Binhai New Area Health Commissions [KJ20071 & 2019BWKY010]; Tianjin Science and Technology Project [18PTZWHZ00060]; TEDA International

Cardiovascular Hospital [2021-TD-006 & 2021-ZX-002 & 2019-TD-013]; Tianjin Key Medical Discipline (Specialty) Construction Project [TJYXZDXK-019A].

Availability of data and materials

The authors are not able to share the clinical data due to full anonymisation of the data is very difficult. The reference gene sequence in our study was based on chromosome 14 (<https://www.ncbi.nlm.nih.gov/nucore/301898188>). All sequencing data used to support the findings of this study are available from <https://doi.org/10.6084/m9.figshare.20110322.v1> and the corresponding author upon request.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the principles of the Declaration of Helsinki. Approval for the study was obtained from the ethics committee of TEDA International Cardiovascular Hospital, China (No. 0715-4, 2021). Written informed consent was obtained from the participants or the parents of minors.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Cardiovascular Surgery, The Institute of Cardiovascular Diseases, TEDA International Cardiovascular Hospital, Tianjin University & Chinese Academy of Medical Sciences, No. 61, the 3rd Ave., Tianjin 300457, China.

²Clinical School of Cardiovascular Disease, Tianjin Medical University, Tianjin, China.

Received: 14 June 2022 Accepted: 21 September 2022

Published online: 08 October 2022

References

- Chowdhury D, Johnson JN, Baker-Smith CM, Jaquiss R, Mahendran AK, Curren V, et al. Health care policy and congenital heart disease: 2020 focus on our 2030 future. *J Am Heart Assoc.* 2021;10(20):e020605.
- Samad T, Wu SM. Single cell RNA sequencing approaches to cardiac development and congenital heart disease. *Semin Cell Dev Biol.* 2021;118:129–35.
- Simmons W, Lin S, Luben TJ, Sheridan SC, Langlois PH, Shaw GM, et al. Modeling complex effects of exposure to particulate matter and extreme heat during pregnancy on congenital heart defects: a US population-based case-control study in the National Birth Defects Prevention Study. *Sci Total Environ.* 2022;808:152150.
- GBD 2017 Congenital Heart Disease Collaborators. Global, regional, and national burden of congenital heart disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Child Adolesc Health.* 2020;4(3):185–200.
- Stout KK, Daniels CJ, Aboulhosn JA, Bozkurt B, Broberg CS, Colman JM, et al. AHA/ACC guideline for the management of adults with congenital heart disease: executive summary: a report of the American college of cardiology/American heart association task force on clinical practice guidelines. *Circulation.* 2019;139(14):e637–97.
- Hou HT, Chen HX, Wang XL, Yuan C, Yang Q, Liu ZG, et al. Genetic characterisation of 22q11.2 variations and prevalence in patients with congenital heart disease. *Arch Dis Child.* 2020;105(4):367–74.
- Zheng SQ, Chen HX, Liu XC, Yang Q, He GW. Genetic analysis of the CITED2 gene promoter in isolated and sporadic congenital ventricular septal defects. *J Cell Mol Med.* 2021;25(4):2254–61.
- Penny DJ, Vick GW 3rd. Ventricular septal defect. *Lancet.* 2011;377(9771):1103–12.
- Zhao QM, Liu F, Wu L, Ma XJ, Niu C, Huang GY. Prevalence of congenital heart disease at live birth in China. *J Pediatr.* 2019;204:53–8.
- Miyake T. A review of isolated muscular ventricular septal defect. *World J Pediatr.* 2020;16(2):120–8.
- Lu SH, Lee KZ, Hsu PW, Su LY, Yeh YC, Pan CY, et al. Alternative splicing mediated by RNA-binding protein RBM24 facilitates cardiac myofibrillogenesis in a differentiation stage-specific manner. *Circ Res.* 2022;130(1):112–29.
- Mahdavi V, Chambers AP, Nadal-Ginard B. Cardiac alpha- and beta-myosin heavy chain genes are organized in tandem. *Proc Natl Acad Sci USA.* 1984;81(9):2626–30.
- Arrington CB, Bleyl SB, Matsunami N, Bonnell GD, Otterud BE, Nielsen DC, et al. Exome analysis of a family with pleiotropic congenital heart disease. *Circ Cardiovasc Genet.* 2012;5(2):175–82.
- Carniel E, Taylor MR, Sinagra G, Di Lenarda A, Ku L, Fain PR, et al. Alpha-myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy. *Circulation.* 2005;112(1):54–9.
- Ching YH, Ghosh TK, Cross SJ, Packham EA, Honeyman L, Loughna S, et al. Mutation in myosin heavy chain 6 causes atrial septal defect. *Nat Genet.* 2005;37(4):423–8.
- Thorolfsson RB, Sveinbjornsson G, Aegisdottir HM, Benonisdottir S, Stefansdottir L, Ivarsdottir EV, et al. Genetic insight into sick sinus syndrome. *Eur Heart J.* 2021;42(20):1959–71.
- Huang S, Wu Y, Chen S, Yang Y, Wang Y, Wang H, et al. Novel insertion mutation (Arg1822_Glu1823dup) in MYH6 coiled-coil domain causing familial atrial septal defect. *Eur J Med Genet.* 2021;64(11):104314.
- Theis JL, Hu JJ, Sundsbak RS, Evans JM, Bamlet WR, Qureshi MY, et al. Genetic association between hypoplastic left heart syndrome and cardiomyopathies. *Circ Genom Precis Med.* 2021;14(1):e003126.
- Bartosovic M, Kabbe M, Castelo-Branco G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat Biotechnol.* 2021;39(7):825–35.
- Kadiyala P, Carney SV, Gauss JC, Garcia-Fabiani MB, Haase S, Alghamri MS, et al. Inhibition of 2-hydroxyglutarate elicits metabolic reprogramming and mutant IDH1 glioma immunity in mice. *J Clin Invest.* 2021;131(4):139542.
- Marino N, German R, Podicheti R, Rusch DB, Rockey P, Huang J, et al. Aberrant epigenetic and transcriptional events associated with breast cancer risk. *Clin Epigenetics.* 2022;14(1):21.
- Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Berhanu Lemma R, Turchi L, Blanc-Mathieu R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 2022;50(D1):D165–73.
- Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* 2004;32:D91–4.
- Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 2021;49(D1):D605–12.
- Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape automation: empowering workflow-based network analysis. *Genome Biol.* 2019;20(1):185.
- Hellman LM, Fried MG. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc.* 2007;2(8):1849–61.
- Holm H, Gudbjartsson DF, Sulem P, Masson G, Helgadóttir HT, Zanon C, et al. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet.* 2011;43(4):316–20.
- Abu-Daya A, Sater AK, Wells DE, Mohun TJ, Zimmerman LB. Absence of heartbeat in the *Xenopus tropicalis* mutation muzak is caused by a nonsense mutation in cardiac myosin myh6. *Dev Biol.* 2009;336(1):20–9.
- England J, Loughna S. Heavy and light roles: myosin in the morphogenesis of the heart. *Cell Mol Life Sci.* 2013;70(7):1221–39.
- van der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol.* 2011;58(21):2241–7.
- Zhang J, Zhang Y, He X, Wang S, Pang S, Yan B. TFEB gene promoter variants effect on gene expression in acute myocardial infarction. *Front Cell Dev Biol.* 2021;9:630279.

32. Zheng SQ, Chen HX, Liu XC, Yang Q, He GW. Identification of variants of ISL1 gene promoter and cellular functions in isolated ventricular septal defects. *Am J Physiol Cell Physiol*. 2021;321(3):443–52.
33. Wan C, Borgeson B, Phanse S, Tu F, Drew K, Clark G, et al. Panorama of ancient metazoan macromolecular complexes. *Nature*. 2015;525(7569):339–44.
34. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al. Architecture of the human interactome defines protein communities and disease networks. *Nature*. 2017;545(7655):505–9.
35. Sadayappan S, Osinska H, Klevitsky R, Lorenz JN, Sargent M, Molkenin JD, et al. Cardiac myosin binding protein C phosphorylation is cardioprotective. *Proc Natl Acad Sci USA*. 2006;103(45):16918–23.
36. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature*. 2020;578(7793):82–93.
37. Feng Y, Li M, Wang S, Cong W, Hu G, Song Y, et al. Paired box 6 inhibits cardiac fibroblast differentiation. *Biochem Biophys Res Commun*. 2020;528(3):561–6.
38. Lilly AJ, Lacaud G, Kouskoff V. SOXF transcription factors in cardiovascular development. *Semin Cell Dev Biol*. 2017;63:50–7.
39. Yu W, Wang M, Zhang H, Quan Y, Zhang Y. Expression and functional analysis of storage protein 2 in the silkworm, *bombyx mori*. *Int J Genomics*. 2013;2013:145450.
40. Chen HX, Zhang X, Hou HT, Wang J, Yang Q, Wang XL, et al. Identification of a novel and functional mutation in the TBX5 gene in a patient by screening from 354 patients with isolated ventricular septal defect. *Eur J Med Genet*. 2017;60(7):385–90.
41. Naiche LA, Harrelson Z, Kelly RG, Papaioannou VE. T-box genes in vertebrate development. *Annu Rev Genet*. 2005;39:219–39.
42. Guo Z, Geng M, Qin L, Hao B, Liao S. Epicardium-derived Tbx18(+) CDCs transplantation improve heart function in infarcted mice. *Front Cardiovasc Med*. 2021;8:744353.
43. Hu Z, Shi Y, Mo X, Xu J, Zhao B, Lin Y, et al. A genome-wide association study identifies two risk loci for congenital heart malformations in Han Chinese populations. *Nat Genet*. 2013;45(7):818–21.
44. Krause A, Zacharias W, Camarata T, Linkhart B, Law E, Lischke A, et al. Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Dev Biol*. 2004;273(1):106–20.
45. Maeda J, Yamagishi H, McAnally J, Yamagishi C, Srivastava D. Tbx1 is regulated by forkhead proteins in the secondary heart field. *Dev Dyn*. 2006;235(3):701–10.
46. Sadahiro T, Isomi M, Muraoka N, Kojima H, Haginiwa S, Kurotsu S, et al. Tbx6 induces nascent mesoderm from pluripotent stem cells and temporally controls cardiac versus somite lineage diversification. *Cell Stem Cell*. 2018;23(3):382–95.e5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

