ORIGINAL ARTICLE

SINGLE-NUCLEOTIDE POLYMORPHISMS IN EXONIC AND PROMOTER REGIONS OF TRANSCRIPTION FACTORS OF SECOND HEART FIELD ASSOCIATED WITH SPORADIC CONGENITAL CARDIAC ANOMALIES

Wang E[1,*], Fan X[3,*], Nie Y[2], Zheng Z[1], Hu S[1,2,#]

*These authors contribute equally to this article

#**Corresponding Author:** Shengshou Hu M. D., Cardiac Surgery Department, Fuwai Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Xicheng District, Beijing, 100037, China, Tel & Fax: 86-010-88322325 E-mail: shengshouh@sina.cn

ABSTRACT

Multiple second heart field (SHF) transcription factors are involved in cardiac development. In this article we evaluate the relationship between SHF transcription factor polymorphisms and congenital heart disease (CHD). Ten polymorphisms were used for genotyping, and three of these were used for the luciferase assay. The risk of CHD was increased 4.31 times and 1.54 times in the C allele of *GATA5*: rs6061243 G>C and G allele of *TBX20*: rs336283 A>G, respectively. The minor alleles of *SMYD1*: rs1542088 T>G, *MEF2C*: rs80043958 A>G and *GATA5*: rs6587239 T>C increased the risk of the simple types of CHD. The minor alleles of *GATA5*: rs41305803 G>A and *MEF2C*: rs304154 A>G increased the risk of tetralogy of Fallot (TOF). The minor alleles of *TBX20*: rs336284 A>G and *SMYD1*: rs88387557 T>G only increased the risk of a single ventricle (SV). Luciferase assays revealed that the minor alleles of rs304154 and rs336284 decreased the transcriptional levels of *MEF2C* and *TBX20,* respectively (p<0.01). When combined with *HLTF*, the G promoter showed a higher expression level than the A promoter in rs80043958 (p<0.01). Our findings suggest that minor alleles of SNPs in the exonic and promoter regions of transcription factors in the SHF can increase the risks of sporadic CHD.

Key words: Polymorphisms; Transcription factor; Congenital cardiac anomaly

INTRODUCTION

Congenital cardiac anomalies are the main cause of infant death and the most common birth defect worldwide (1). Patients with sporadic congenital heart disease (CHD) account for approximately 80% of CHD patients (2). Depending on the different anatomic or pathophysiological changes, CHD can be divided into 21 different forms, including simple common forms of CHD (1) and moderate and severe forms of CHD (1). Although the incidence rate of CHD is high and its clinical symptoms are obvious, its etiology is still unclear in most patients (3).

Cardiac progenitor cells from the second heart field (SHF) participate in the development of linear cardiac tubes when the cardiac tube becomes the four-chambered heart (4). More than 10 transcription factors, including *GATA5* (GenBank accession no. NM_080473), *MEF2C* (GenBank accession no. NM_002397), *SYMD1* (GenBank accession no. NM_198274), and *TBX20* (GenBank accession no. NM_001166220), contribute to SHF development by controlling the proliferation and differentiation of cardiac progenitor cells (5). The knockout of these genes could lead to different types of CHD in mice (6-9). Many exonic mutations of SHF transcription factors (*GATA5*, *MEF2C*, *SYMD1*, and *TBX20*) are related to CHD in humans (10-13). However, the underlying genetic pathogenesis of CHD remains unclear. In this study, we reveal that minor alleles of ten exonic and promoter single nucleotide

[1] Cardiac Surgery Department, Fuwai Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, 100037, China

[2] State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, 100037, China

[3] Clinical Laboratory Center, Beijing An Zhen Hospital, Capital Medical University, Beijing Institute of Heart, Lung and Blood Vessel Diseases Beijing, 100029, China

polymorphisms (SNPs) located in SHF transcription factors increase sporadic CHD risk.

MATERIALS AND METHODS

Patient information

From January 13, 2012 to May 5, 2012, a total of 383 patients were enrolled in this study. These patients were suffering from sporadic CHD and were scheduled for surgery in our hospital (Table 1). The average age of the patients was 1 year old (3 months - 9 years). Patients were classified into simple CHD [ventricular septal defects (VSD), atrial septal defects (ASD), and patent ductus arteriosus (PDA)] (33%), right ventricular outflow tract obstruction (RVOTO) [tetralogy of Fallot (TOF), pulmonary atresia (PA), and pulmonary stenosis (PS)] (58%), and single ventricle (SV) (9%) (Table 1). A total of 383 healthy children were also recruited from our hospital as a control group. No significant differences were observed in age or sex between the CHD patients and control subjects (Table S1, S2). The diagnosis and inclusion and exclusion criteria for the subjects are described in the "Methods section" of the supplementary material. The study complied with the 1964 Declaration of Helsinki and its subsequent amendments and was approved by the Medical Ethics Committee of Fuwai Hospital. All patients or their legal guardians signed informed consent forms.

DNA extraction and genotyping

In all subjects, the extraction of genomic DNA from leukocytes was performed with a Wizard® Genomic DNA Purification Kit (Promega, WI, USA). Ten SNPs in ex-

Table 1. Patients' Information

	Number(%)
No. of patients (n)	383
Age (years)	2.22±1.99
Male gender(n, %)	216(56%)
Clinical diagnosis (n, %)	
simple CHD	125(33%)
VSD	94(25%)
ASD	19(5%)
PDA	12(3%)
ROVTO	223(58%)
TOF	107(28%)
PA or PS with VSD	67(17%)
PA or PS with IVS	49(13%)
SV	35(9%)

PA: Pulmonary Atresia; PS: Pulmonary Stenosis; IVS: Intact Ventricular Septum; RVOTO: Right Ventricle Outflow Tract Obstruction; ASD: Atrial Septal Defect; VSD: Ventricular Septal Defect; PDA: Patent Ductus Arteriosus; TOF:Tetralogy Of Fallot; SV: Single Ventricle.
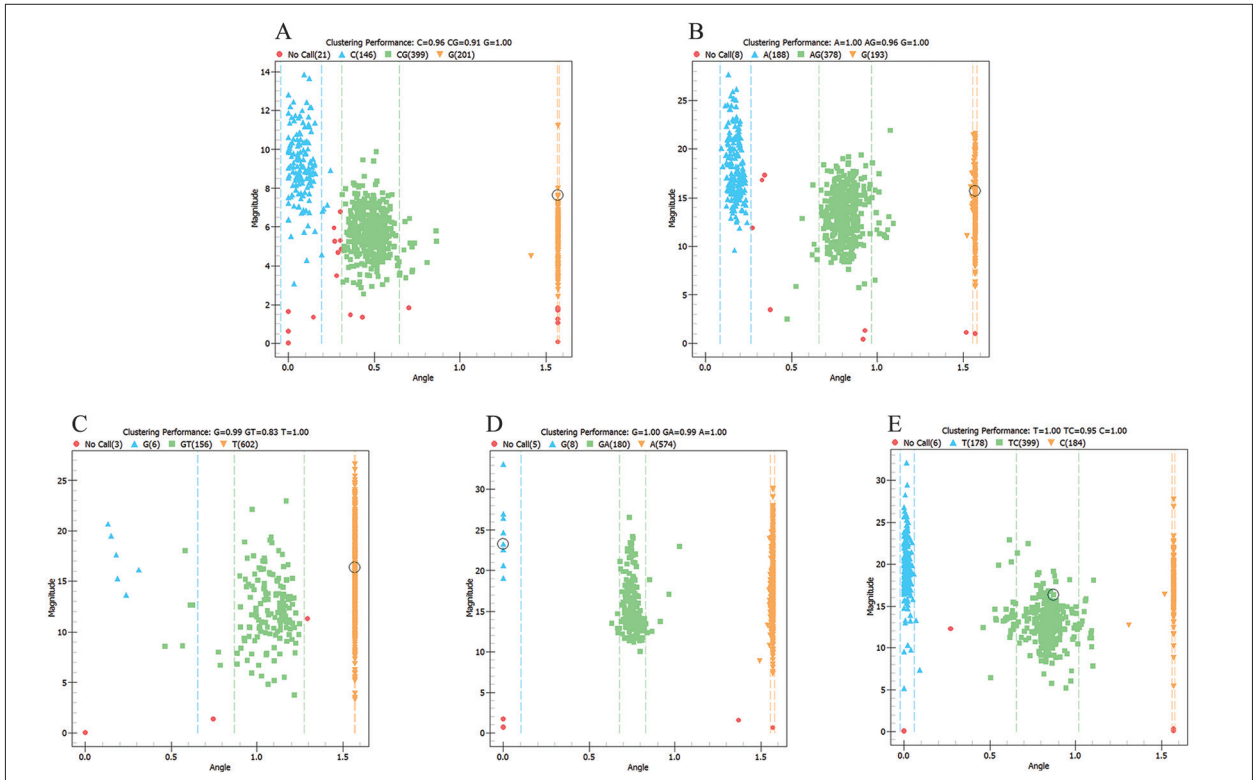
onic and promoter regions from 4 genes (*GATA5*, *SMYD1*, *TBX20*, and *MEF2C*) (from unpublished sequencing data for CHD) were genotyped by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) in both CHD patients and control subjects (Figure 1, 2). The steps of MALDI-TOF-MS included multiplex polymerase chain reaction (PCR), amplification, shrimp alkaline phosphatase digestion, IPLEX primer extension, resin cleaning, MALDI-TOF-MS analysis, and data analysis (14). Analyses were repeated in 10% of randomly selected samples for quality control.

Plasmids, site-directed mutagenesis, cell transfection, and luciferase assays

MEF2C: rs80043958 A>G, *MEF2C*: rs304154 A>G, and *TBX20*: rs336284 A>G were all located in promoter elements. The promoter fragments of *MEF2C* containing the A allele of rs80043958 or rs304154 and the *TBX20* promoter fragment containing the A allele of rs336284 were amplified from genomic DNA. The PCR products were subcloned into the KpnI/XhoI restriction sites of the GV238-basic vector (GeneChem, Shanghai, China). Transcription factors included *ZFX* (GenBank accession no. NM_001330327), *CEBPA* (GenBank accession no. NM_001287424), *HLTF* (GenBank accession no. NM_003071), *FOXC1* (GenBank accession no. NM_001453), and *GATA1* (GenBank accession no. NM_002049) were also amplified and subcloned into the GV141-basic vector. Plasmids carrying the corresponding G allele were generated by site-directed mutagenesis with the MutanBEST kit (Takara, Berkeley, CA, USA) to ensure a uniform backbone sequence. All recombinant clones were verified by DNA sequencing. The human embryonic kidney cell line HEK 293T (4×10^5) was seeded in 24 well culture plates. After 24 h, HEK 293T cells were transfected with 1.0 µg of the wild-type promoter or mutant promoter and the corresponding transcription factors, according to the manufacturer's instructions. After an additional 24 h of culture, the transfected cells were assayed for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega). There were eight experimental groups for rs80043958 and rs304154. For rs336284, there were four experimental groups. Each luciferase assay was performed in triplicate.

Statistical analyses

The means ± standard deviations (SD) were used for the continuous variables. Continuous variables were compared between the two groups by Student's t test. Pearson's χ^2 test or Fisher's exact test was used to compare the categorical variables between the two groups. The odds ratios

Figure 1. Genotyping of five exonic and promoter SNPs located in *GATA5*, *TBX20*, *SMYD1* and *MEF2C*.
Genotyping of (A) *GATA5:* rs6061243 G>C, (B) *TBX20*: rs336283 A>G, (C) *SMYD1*: rs1542088 T>G, (D) *MEF2C*: rs80043958 A>G and (E) *GATA5*: rs6587239 T>C SNPs by MALDI-TOF-MS.
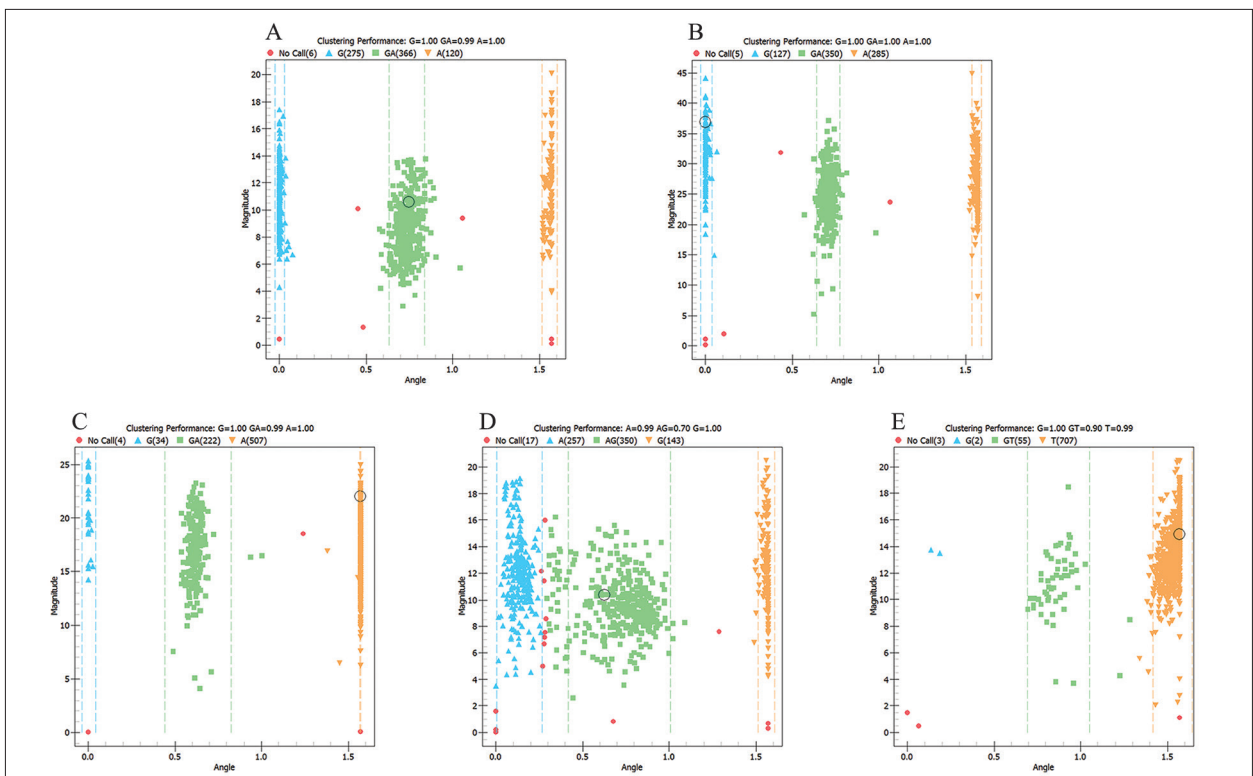

Figure 2. Genotyping of the other five SNPs located in *GATA5*, *TBX20*, *SMYD1* and *MEF2C* exons and promoters.
Genotyping of (A) *GATA5*: rs41305803 G>A, (B) *MEF2C*: rs304154 A>G, (C) *SMYD1*: rs2919881 A>G, (D) *TBX20*: rs336284 A>G and (E) *SMYD1*: rs88387557 T>G SNPs by MALDI-TOF-MS.

(ORs) and the corresponding 95% confidence intervals (CIs) were estimated for the risk of CHD. Differences were considered significant if p<0.05. The statistical analyses were performed using the SPSS version 17.0 software package (SPSS, Inc., Chicago, IL, USA).

RESULTS

Ten SNPs located in exon and promoter regions were genotyped

There were seven SNPs in exons and three SNPs in promoters (Table S3). Among the seven exonic SNPs, *SMYD1*: rs88387557 T>G was a nonsynonymous mutation, while the remaining six SNPs were all synonymous mutations (Table S3). The bioinformatics analysis suggested that the other three SNPs located in promoters caused promoter loss, which might influence *MEF2C* and *TBX20* mRNA transcription (Table S3). The polymorphism rates of genotyping were 98-100%. The genotype frequencies of the controls were in accordance with Hardy-Weinberg equilibrium (p>0.05) (Table S4).

The risk of CHD was increased by minor alleles of rs6061243 and rs336283

The minor alleles of two SNPs, including the C allele of *GATA5*: rs6061243 G>C and the G allele of *TBX20*: rs336283 A>G, significantly increased the risk of CHD.

Subjects carrying *GATA5*: rs6061243 GC had a 4.31-fold increased risk of CHD (OR=4.31, 95% CI 3.03-6.13, p=1.03×10^{-16}); additionally, compared with patients with wild-type *TBX20*: rs336283 AA, AG increased the risk of CHD by 1.54-fold (OR=1.54; 95% CI 1.08-2.19, p<0.05) (Table 2). The subjects with *TBX20*: rs336283 GG were associated with a 1.91-fold increased risk of CHD in comparison with patients carrying AA (OR=1.91; 95% CI 1.27-2.87, p=0.002). The remaining minor alleles of the 8 polymorphisms were not associated with CHD risk (Table 2).

Subgroup analyses were then implemented to assess the impacts of *GATA5*: rs6061243 G>C and *TBX20*: rs336283 A>G polymorphisms on CHD subtypes. The minor C allele of *GATA5* rs6061243 G>C increased the risk of different CHD subtypes (p<0.05) (Table S5). However, GG in *TBX20*: rs336283 A>G increased the risk of various CHD subtypes but not simple CHD (p<0.05) (Table S6).

The minor alleles of *SMYD1*: rs1542088 T>G, *MEF2C*: rs80043958 A>G and *GATA5*: rs6587239 T>C increase the risk of simple CHD

The minor alleles in *SMYD1*: rs1542088 T>G, *MEF2C*: rs80043958 A>G and *GATA5*: rs6587239 T>C had no associations with the risk of CHD but were associated with the risk of simple CHD, including VSD, ASD, and PDA. Additionally, these three SNPs did not change the RVOTO or SV risk. For *MEF2C*: rs80043958 A>G, subjects carrying GG+GA had an increased simple CHD risk in comparison with patients carrying AA homozy-

Table 2. Main effects of SNPs on CHD risk

SNP	Homozygotes (for common alleles)		Heterozygotes				Homozygotes (for rarer alleles)			
	Case	Control	Case	Control	OR (95%CI)	p value	Case	Control	OR (95%CI)	p value
GATA5: rs6061243 G>C	103	182	200	82	**4.31 (3.03-6.13)**	**1.03×10^{-16}**	70	119	1.04 (0.71-1.52)	0.84
TBX20: rs336283 A>G	76	112	193	185	**1.54 (1.08-2.19)**	**0.017**	109	84	**1.91 (1.27-2.87)**	**0.002**
SMYD1: rs1542088 T>G	292	310	83	73	1.21 (0.85-1.72)	0.3	6	0	–	–
MEF2C: rs80043958 A>G	280	294	96	83	1.21 (0.87-1.70)	0.26	5	3	1.75 (0.41-7.39)	0.44
GATA5: rs6587239 T>C	88	90	196	203	0.99 (0.69-1.41)	0.94	96	88	1.12 (0.74-1.69)	0.6
GATA5: rs41305803 G>A	128	147	188	178	1.21 (0.89-1.66)	0.23	62	58	1.23 (0.78-1.89)	0.35
MEF2C: rs304154 A>G	134	151	188	162	1.31 (0.96-1.79)	0.09	58	69	0.95 (0.62-1.44)	0.8
SMYD1: rs2919881 A>G	257	250	106	116	0.89 (0.65-1.22)	0.47	18	16	1.09 (0.55-2.19)	0.8
TBX20: rs336284 A>G	126	131	169	181	0.97 (0.70-1.34)	0.86	77	66	1.21 (0.81-1.83)	0.36
SMYD1: rs88387557 T>G	350	357	31	24	1.32 (0.76-2.29)	0.33	1	1	1.02 (0.06-16.37)	1

gotes (OR=1.59, 95% CI 1.02-2.48, p=0.04) (Table 3). For *SMYD1*: rs1542088 T>G, subjects carrying GT were associated with an increased risk of simple CHD compared with TT subjects (OR=1.62, 95% CI 1.01-2.60, p=0.043) (Table S7). GG+GT increased the risk of simple CHD compared with TT (OR = 1.72, 95% CI 1.08-2.73, p=0.021) (Table S7). *GATA5*: rs6587239 CC was associated with an increased risk of simple CHD compared with TT+TC (OR=1.59, 95% CI 1.02-2.48, p=0.042) (Table S8).

The minor A allele of *GATA5*: rs41305803 G>A and the minor G allele of *MEF2C*: rs304154 A>G increase the risk of RVOTO or TOF

The minor G allele of *MEF2C*: rs304154 A>G only increased TOF risk. GA increased the TOF risk by 1.67-fold in comparison with AA (OR=1.67, 95% CI = 1.03-2.69, p = 0.036) (Table 4). The minor A allele of *GATA5*: rs41305803 G>A increased the risk of RVOTO. The RVOTO group was then divided into three subgroups: TOF, PA or PS with VSD and PA or PS with IVS. The minor A allele of *GATA5*: rs41305803 G>A only increased the risk of TOF (Table S9).

The minor alleles of *TBX20*: rs336284 A>G, *SMYD1*: rs2919881 A>G, and *SMYD1*: rs88387557 T>G increase the risk of other CHD types

Among patients harboring *TBX20*: rs336284 A>G, GG subjects were significantly associated with an increased risk of SV compared with AA subjects (OR =2.26, 95% CI 1.05-4.86, p=0.033) (Table 5). However, among patients harboring *SMYD1*: rs2919881 A>G, the risk of PA or PS with IVS was increased significantly in GG subjects compared with AA patients (OR =3.19, 95% CI 1.19-8.59, p= 0.039) (Table S10). For *SMYD1*: rs88387557 T>G, the SV risk was increased by 2.66-fold in the G allele compared with the T allele (OR=2.66, 95% CI 1.06-6.70, p=0.03) (Table S11).

Luciferase assays of *MEF2C*: rs80043958 A>G, *MEF2C*: rs304154 A>G, and *TBX20*: rs336284 A>G.

For rs80043958, the G allele plasmid showed non-significant luciferase expression compared with the A allele counterparts in HEK 293T cells (p>0.05) (Figure 3A). When combined with *HLTF*, the G promoter showed a higher expression level than the A promoter (p<0.01) (Figure 3A). No such increase occurred when *CEBPA* or *CEBPA+HLTF* were added to the *MEF2C* promoter (Figure 3A).

Table 3: *MEF2C*: rs80043958 A>G

Genotype	simple CHD (n=125)		Control (n=381)		OR (95%CI)	p value
	No	(%)	No.	(%)		
AA	85	68	294	77	1	
GA	37	30	84	22	1.52(0.97-2.40)	0.07
GG	3	2	3	1	3.46(0.69-17.45)	0.11
GG+GA	40	32	87	23	**1.59(1.02-2.48)**	**0.04**
AA+GA	122	98	378	99	1	
GG	3	10	3	4	3.10(0.62-15.55)	0.15
A allele	207	83	672	88	1	
G allele	43	17	90	12	**1.55(1.05-2.30)**	**0.029**

CHD: congenital heart disease

Table 4: *MEF2C*: rs304154 A>G

Genotype	TOF (n=105)		Control (n=382)		OR (95%CI)	p value
	No	(%)	No.	(%)		
AA	33	31	151	40	1	
GA	59	56	162	42	1.67(1.03-2.69)	0.036
GG	13	12	69	18	0.86(0.43-1.74)	0.68
GG+GA	72	69	231	60	1.43(0.90-2.26)	0.13
AA+GA	92	88	313	82	1	
GG	13	10	69	4	0.64(0.34-1.21)	0.17
A allele	125	60	464	61	1	
G allele	85	40	300	39	1.05(0.77-1.44)	0.75

TOF: tetralogy of Fallot

Table 5: *TBX20*: rs336284 A>G

Genotype	SV (n=34)		Control (n=378)		OR (95%CI)	p value
	No	(%)	No.	(%)		
AA	9	26	131	35	1	
GA	14	41	181	48	1.13(0.47-2.68)	0.79
GG	11	32	66	17	2.42(0.96-6.14)	0.06
GG+GA	25	74	247	65	1.47(0.67-3.25)	0.33
AA+GA	23	68	312	83	1	
GG	11	10	66	4	2.26(1.05-4.86)	0.033
A allele	32	47	443	59	1	
G allele	36	53	313	41	1.59(0.97-2.62)	0.07

SV: single ventricle

For rs304154, the G promoter displayed a significantly lower luciferase expression than the A promoter (p<0.01) (Figure 3B). When the *MEF2C* promoter was combined with *GATA1*, *FOXC1*, or *GATA1+FOXC1*, the two groups still exhibited a significant difference (p<0.01) (Figure 3B).

For rs336284, the G promoter exhibited a significantly lower level of luciferase expression than the A promoter (p<0.01) (Figure 3C). The promoter in the (G)+*ZFX* group showed a lower expression level (p<0.01).
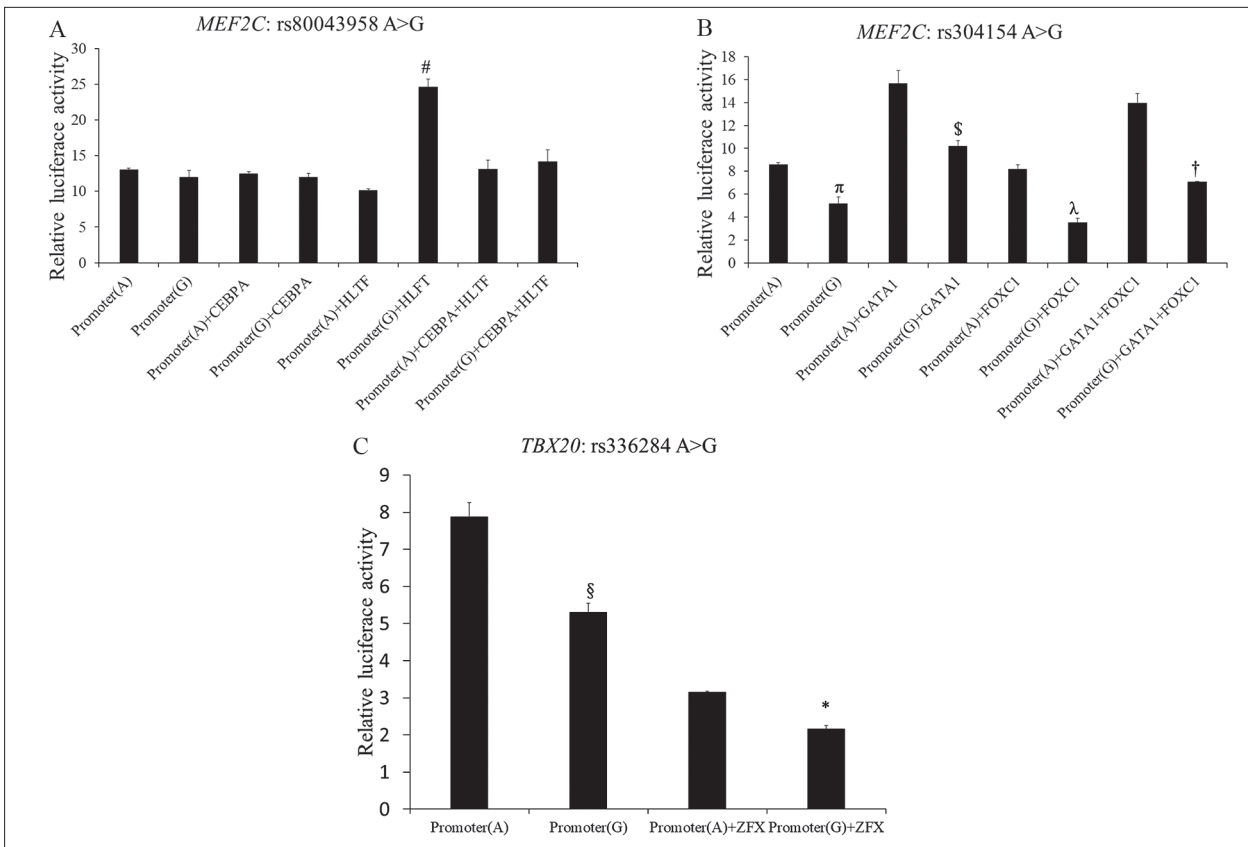
Figure 3. Luciferase assays of *MEF2C*: rs80043958 A>G, *MEF2C*: rs304154 A>G, and *TBX20*: rs336284 A>G

Figure 3A, For rs80043958, the plasmid containing the G allele displayed nonsignificant luciferase expression compared with the wild-type A allele (p>0.05). When combined with HLTF, the (G) promoter showed a higher expression level than the (A) promoter (#, p<0.01);

Figure 3B, For rs304154, the (G) promoter displayed significant lower luciferase expression than the (A) promoter (π, p<0.01). When the *MEF2C* promoter was combined with *GATA1*, *FOXC1*, or *GATA1+FOXC1*, the difference was still significant between the two groups ($, λ, †, p<0.01);

Figure 3C, For rs336284, the (G) promoter exhibited a significantly lower luciferase expression level than the (A) promoter (§, p<0.01). The promoter (G)+ZFX group also showed a lower expression level (*, p<0.01).

DISCUSSION

SNPs located in exon or promoter regions play important roles in CHD development. Our previous study revealed that polymorphisms located in the exons and promoters of growth factors influence the risk of CHD (15). We then revealed that the minor alleles in the intronic SNPs in the transcription factors regulating the second heart field increased the risk of CHD (14). In this study, we evaluated the associations between polymorphisms in the exons and promoters of the second heart field and CHD.

Mutations in *SMYD1* (16), *MEF2C* (11), *GATA5* (17), and *TBX20* (10) were found to be critical for the transcriptional regulation of gene expression for CHD development.

Mutations in the exons of these genes can inactivate an allele or cause gene dysfunction by interfering with DNA interactions. In this study, we found ten SNPs located in the exons and promoters of *SMYD1*, *MEF2C*, *GATA5*, and *TBX20* which regulate SHF development and were significantly associated with CHD. The minor alleles of two SNPs, *GATA5*: rs6061243 G>C and *TBX20*: rs336283 A>G, were significantly associated with an increased risk of CHD. Additionally, the minor alleles of the remaining SNPs increased the risk of different CHD types.

SHF regulation involves numerous signaling and transcriptional cascades. The ISL1-GATA-MEF2C pathway plays an important and central role in the transcription factor network of SHF. When heart looping occurs, *Mef2c* transcripts are robustly expressed in the outflow tract and right ventricle and are less abundant in the left ventricle and atria (18). Mouse embryos lacking *Mef2c* exhibit severe defects of the outflow tract and hypoplasia of the right ventricle (19). These observations imply a crucial role for *MEF2C* in the transcription factor networks regulating myoblast

differentiation in SHFs. There were two SNPs located in the promoter region of *MEF2C* that were associated with CHD in this study (15). The minor G allele of *MEF2C*: rs80043958 A>G exhibited an increased risk of VSD, ASD, and PDA, while the G allele of *MEF2C*: rs304154 A>G only increased TOF risk. The bioinformatic analysis suggested that both SNPs cause promoter loss in *MEF2C*, which might disrupt binding with other transcription factors and influence *MEF2C* transcription. Subsequent luciferase assays showed that the minor G allele of rs304154 could either decrease the transcription level of *MEF2C* alone or influence its transcription level in combination with *FOXC1*, *GATA1*, or *FOXC1+GATA1*. Although we did not find that the G allele of rs80043958 influenced the transcriptional level of *MEF2C* when the *HLTF* interacted with a promoter containing rs80043958. The minor G allele increased the transcription level of *MEF2C*.

Epigenetic factors may also be involved in cardiac morphogenesis. The disturbance of the chromatin remodeling protein *Smyd1* in mice results in a phenotype with a decreased right ventricle and dysplasia of the left ventricle (20). There have been few reports of *SMYD1* mutations associated with CHD. One study revealed that one de novo exonic mutation in *SMYD1* was associated with hypertrophic cardiomyopathy (12). In the present study, we found that both identified SNPs of *SMYD1* were located in the exon. The minor G allele of *SMYD1*: rs2919881 A>G (p.I253I) increased the risk of PA or PS with IVS. The minor G allele of *SMYD1*: rs1542088 T>G (p.R131R) increased the risk of simple CHD. Both SNPs are synonymous variants. There is a SET domain in *Smyd1* that has methyltransferase activity and histone deacetylase (HDAC) activity, which can represses genes (20). p.I253I is located in the ET domain of *SMYD1*, which is a pivotal SET domain that functions as the primary catalytic domain. p.R131R is located in the SET-I domain, which contributes to the binding of cofactors with substrates and protein stability (21). Therefore, we could infer that the p.I253I variant might influence methyltransferase activity, while p.R131R could interfere with cofactor and substrate binding with *SMYD1*.

Three *GATA* factors, *GATA* 4, 5, and 6, were expressed in the heart in a partially overlapping way. *GATA* transcription factors are key regulators of cardiac development. In contrast to *GATA* 4 and 6, *GATA* 5 expression is more restricted to endocardial cushions in the outflow tract during cardiac development. Laforest *et al.* found that *Gata4+/−Gata5+/−* and *Gata5+/−Gata6+/−* double heterozygous mice die in the embryonic or perinatal periods due to dysplasia of the OFT, including double outlet right ventricle (DORV) and VSD. These studies reveal the existence of important genetic interactions between *GATA5* and the other two *Gata* factors in outflow tract morphogenesis (22). Later studies in humans revealed *GATA5* mutations associated with VSD, aortic bicuspid, DORV, and TOF (13, 23-25). In our study, the minor A allele of *GATA5*: rs41305803 G>A (p.D203D) increased the risk of RVOTO, which is an outflow tract malformation. The minor C allele of *GATA5*: rs6587239 T>C (p.K284K) increased the risk of simple CHD. Subjects carrying *GATA5* rs6061243 (p.S327S) GC presented a 4.31-fold increase in the risk of CHD. These results were similar to previous studies and suggest that subtle alterations in the activity of the *GATA5* factor might cause CHD in humans.

TBX20 is critical for heart chamber formation, especially the outflow tract and right ventricle, which are the anterior derivatives of the SHF (26). Targeted disruption of *TBX20* leads to unlooped and severely hypoplastic myocardial tubes in mice (27). Incomplete knockdown of *TBX20* results in severely compromised valve formation, hypoplastic right ventricle, and persistent truncus arteriosus (9). In cardiac development, *TBX20* functions as a dosage-dependent moderator (9). Either loss- or gain-of-function in *TBX20* results in abnormal heart development (28). *TBX20* mutations are associated with VSD, ASD, TOF, DORV, persistent truncus arteriosus (PTA), and adult dilated cardiomyopathy in humans (29, 30). Mutations in either the exon or promoter regions of *TBX20* can lead to CHD (10). This study revealed that *TBX20*: rs336283 A>G in the promoter increased the risk of CHD, while *TBX20*: rs336284 A>G (p.S13S) in an exon increased the risk of SV. Further luciferase assays showed that the G allele of rs336284 decreased the transcription level of *TBX20*, even with the interaction with *ZFX*.

Limitations are still present in this study. First, the small sample size limited the persuasiveness. A large number of CHD patients and healthy controls for genotyping will be included in our next study. Second, the mechanisms of these ten SNPs affecting the risk of CHD still require further investigation. The knock-in mouse model must be built in order to explore the mechanism by which these SNPs affect CHD formation at the genetic and molecular biological levels.

In this study, the associations of exonic and promoter SNPs in SHF transcription factors with an increased risk of CHD were evaluated. This results suggest that SNPs in SHF exon and promoter regions play roles in the pathogenesis of CHD.

Conflict of interest: none

Authors' contributions: Enshi Wang and Shengshou Hu: Conceived of study. Enshi Wang and Xuesong Fan: Performed research. Enshi Wang, Yu Nie, Zhe

REFERENCES

1. Hoffman JI, Kaplan S. The incidence of congenital heart disease. J Am Coll Cardiol. 2002; 39(12): 1890-900.

2. Winkleby M, Sundquist K, Cubbin C. Inequities in CHD incidence and case fatality by neighborhood deprivation. Am J Prev Med. 2007; 32(2): 97-106.

3. Srivastava D. Making or breaking the heart: from lineage determination to morphogenesis. Cell. 2006; 126(6): 1037-48.

4. Rasmussen TL, Ma Y, Park CY, Harriss J, Pierce SA, Dekker JD, et al. Smyd1 facilitates heart development by antagonizing oxidative and ER stress responses. PLoS One. 2015; 10(3): e0121765.

5. Srivastava D, Olson EN. A genetic blueprint for cardiac development. Nature. 2000; 407(6081):221-6.

6. Tsuchihashi T, Maeda J, Shin CH, Ivey KN, Black BL, Olson EN, et al. Hand2 function in second heart field progenitors is essential for cardiogenesis. Developmental Biology. 2011; 351(1): 62-9.

7. Phan D, Rasmussen TL, Nakagawa O, McAnally J, Gottlieb PD, Tucker PW, et al. BOP, a regulator of right ventricular heart development, is a direct transcriptional target of MEF2C in the developing heart. Development. 2005; 132(11): 2669-78.

8. Vong L, Bi W, O'Connor-Halligan KE, Li C, Cserjesi P, Schwarz JJ. MEF2C is required for the normal allocation of cells between the ventricular and sinoatrial precursors of the primary heart field. Developmental Dynamics. 2006; 235(7):1809-21.

9. Takeuchi JK, Mileikovskaia M, Koshiba-Takeuchi K, Heidt AB, Mori AD, Arruda EP, et al. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. Development. 2005; 132(10): 2463-74.

10. Huang RT, Wang J, Xue S, Qiu XB, Shi HY, Li RG, et al. TBX20 loss-of-function mutation responsible for familial tetralogy of Fallot or sporadic persistent truncus arteriosus. Int J Med Sci. 2017; 14(4): 323-32.

11. Lu CX, Wang W, Wang Q, Liu XY, Yang YQ. A Novel MEF2C Loss-of-Function Mutation Associated with Congenital Double Outlet Right Ventricle. Pediatr Cardiol. 2018; 39(4): 794-804.

12. Fan LL, Ding DB, Huang H, Chen YQ, Jin JY, Xia K, et al. A de novo mutation of SMYD1 (p.F272L) is responsible for hypertrophic cardiomyopathy in a Chinese patient. Clin Chem Lab Med. 2019; 57(4): 532-9.

13. Kassab K, Hariri H, Gharibeh L, Fahed AC, Zein M, El-Rassy I, et al. GATA5 mutation homozygosity linked to a double outlet right ventricle phenotype in a Lebanese patient. Mol Genet Genomic Med. 2016; 4(2): 160-71.

14. Wang E, Nie Y, Fan X, Zheng Z, Hu S. Intronic Polymorphisms in Gene of Second Heart Field as Risk Factors for Human Congenital Heart Disease in a Chinese Population. DNA Cell Biol. 2019; 38(6): 521-31.

15. Wang E, Wang Z, Liu S, Gu H, Gong D, Hua K, et al. Polymorphisms of VEGF, TGFbeta1, TGFbetaR2 and conotruncal heart defects in a Chinese population. Mol Biol Rep. 2014; 41(3): 1763-70.

16. Topf A, Griffin HR, Glen E, Soemedi R, Brown DL, Hall D, et al. Functionally significant, rare transcription factor variants in tetralogy of Fallot. PLoS One. 2014; 9(8): e95453.

17. Hempel M, Casar Tena T, Diehl T, Burczyk MS, Strom TM, Kubisch C, et al. Compound heterozygous GATA5 mutations in a girl with hydrops fetalis, congenital heart defects and genital anomalies. Hum Genet. 2017; 136(3): 339-46.

18. Edmondson DG, Lyons GE, Martin JF, Olson EN. Mef2 gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. Development. 1994; 120(5): 1251-63.

19. Dodou E, Verzi MP, Anderson JP, Xu SM, Black BL. Mef2c is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. Development. 2004; 131(16): 3931-42.

20. Gottlieb PD, Pierce SA, Sims RJ, Yamagishi H, Weihe EK, Harriss JV, et al. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. Nat Genet. 2002; 31: 25-32.

21. Tracy C, Warren JS, Szulik M, Wang L, Garcia J, Makaju A, et al. The Smyd Family of Methyltransferases: Role in Cardiac and Skeletal Muscle Physiology and Pathology. Curr Opin Physiol. 2018; 1: 140-52.

22. Laforest B, Nemer M. GATA5 interacts with GATA4 and GATA6 in outflow tract development. Dev Biol. 2011; 358(2): 368-78.

23. Shi LM, Tao JW, Qiu XB, Wang J, Yuan F, Xu L, *et al.* GATA5 loss-of-function mutations associated with congenital bicuspid aortic valve. Int J Mol Med. 2014; 33(5): 1219-26.

24. Wei D, Bao H, Zhou N, Zheng GF, Liu XY, Yang YQ. GATA5 loss-of-function mutation responsible for the congenital ventriculoseptal defect. Pediatr Cardiol. 2013; 34(3): 504-11.

25. Huang RT, Xue S, Xu YJ, Zhou M, Yang YQ. Somatic GATA5 mutations in sporadic tetralogy of Fallot. Int J Mol Med. 2014; 33(5):1227-35.

26. Cai CL, Liang X, Shi Y, Chu PH, Pfaff SL, Chen J, *et al.* Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. Dev Cell. 2003; 5(6): 877-89.

27. Cai CL, Zhou W, Yang L, Bu L, Qyang Y, Zhang X, *et al.* T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development. 2005; 132(10): 2475-87.

28. Posch MG, Gramlich M, Sunde M, Schmitt KR, Lee SH, Richter S, *et al.* A gain-of-function TBX20 mutation causes congenital atrial septal defects, patent foramen ovale and cardiac valve defects. J Med Genet. 2010; 47(4): 230-5.

29. Pan Y, Geng R, Zhou N, Zheng GF, Zhao H, Wang J, *et al.* TBX20 loss-of-function mutation contributes to double outlet right ventricle. Int J Mol Med. 2015; 35(4): 1058-66.

30. Monroy-Munoz IE, Perez-Hernandez N, Rodriguez-Perez JM, Munoz-Medina JE, Angeles-Martinez J, Garcia-Trejo JJ, *et al.* Novel mutations in the transcriptional activator domain of the human TBX20 in patients with atrial septal defect. Biomed Res Int. 2015; 2015: 718786.