Am J Transl Res 2022;14(11):8407-8415
www.ajtr.org /ISSN:1943-8141/AJTR0145755

Original Article

Associations between *NKX2-5* gene polymorphisms and congenital heart disease in the Chinese Tibetan population

Qiang Ma[1,2], Yingzhong Yang[1], Yongnian Liu[1]

[1]*High Altitude Medical Research Center, Medical College of Qinghai University, 16 Kunlun Road, Xining 810001, Qinghai, China;* [2]*Department of Pathology, Sunshine Union Hospital, Yingqian Road, Weifang 261000, Shandong, China*

Received August 4, 2022; Accepted November 14, 2022; Epub November 15, 2022; Published November 30, 2022

Abstract: Background: The pathogenesis of congenital heart disease (CHD) has not been fully elucidated, and this study considers the interaction between inheritance and the environment as the main cause of CHD. Previous studies have found that the incidence of CHD in the Tibetan plateau population is significantly higher than in low-altitude populations. Numerous reports have confirmed that *NKX2-5* gene mutations can lead to coronary heart disease, but the relationship between *NKX2-5* and Tibetan nationality has not yet been reported. Objective: To explore the relationship between *NKX2-5* gene polymorphisms and CHD in Tibetan people. Methods: Blood samples were collected retrospectively from Tibetan patients diagnosed with CHD as well as healthy Tibetans, and the exons of *NKX2-5* were sequenced. The MassARRAY technique was used to detect and genotype candidate tag single nucleotide polymorphisms (SNPs) in the non-coding regions of *NKX2-5*. Results: Exon sequencing revealed no difference in the coding regions of the *NKX2-5* gene between the CHD and control groups. In the non-coding regions of *NKX2-5*, rs6882776 and rs2546741 differed significantly between the two groups. Strong linkage disequilibrium was found between the selected sites of *NKX2-5*. Conclusions: The *NKX2-5* exons do not associate with CHD in Tibetans. Rs6882776 and rs2546741 in the non-coding regions of *NKX2-5* may protect against CHD in Tibetans. The *NKX2-5* haplotype associated with CHD occurrence in the Tibetan population.

Keywords: Congenital heart disease, *NKX2-5*, single nucleotide polymorphisms, Tibetan, exon, haplotype analysis

Introduction

Congenital heart disease (CHD) is a disease characterized by structural and functional abnormalities in the heart and large blood vessels that occurs during embryonic development. CHD is common and one of the most fatal birth defects [1]. CHD encompasses a broad spectrum of heart malformations ranging from a solitary abnormality (atrial or ventricular septal defect or isolated dysplastic valve) to complex lesions consisting of multiple defects (tetralogy of Fallot or hypoplastic left heart syndrome) [2]. The pathogenesis of CHD has not been fully elucidated; however, it has been suggested that the interaction between inheritance and environment is the cause of CHD [3]. As for environmental factors, the incidence of CHD seems to increase with elevation. The prevalence of CHD in Qinghai province was 0.56%, which was significantly higher than in the low-altitude provinces of China [4]. As for genetic factors, CHD associates with chromosomal abnormalities, polygenic defects, and single-gene mutations [5]. The *NKX2-5* gene is particularly important for the development of the heart. *NKX2-5* is one of the earliest markers of cardiac precursor cell differentiation and is involved in almost all stages of cardiac development, including regulation of the number of cardiac blasts, development of the conduction system, and formation of valves and septa [6]. More than 40 *NKX2-5* gene mutations are associated with CHD [7]. Tibetan people have lived on the high altitude plateau for a long time and they have adapted to a plateau environment. However, early studies have found that the incidence of CHD in Tibetans was signifi-

Table 1. Sequences of primers for the exons of *NKX2-5* gene

Exon	Primer	Sequence	Temperature (°C)	Length (bp)
1	Exon-1-F	5'-CTTGTGCTCAGCGCTACCT-3'	60	550
	Exon-1-R	5'-CTGAGTTTCTTGGGGACGAA-3'		
2	Exon-2-F	5'-AGTGCACTTGGCAGAGTGAG-3'	60	922
	Exon-2-R	5'-CTAGGTCTCCGCAGGAGTGA-3'		

Notes: F, Forward Primer; R, Reverse Primer.

cantly higher than among people who reside at low altitudes [8]. This may be the influence of high-altitude hypoxia and genetics. Numerous reports have confirmed that *NKX2-5* gene mutations lead to the occurrence of CHD, but the relationship between *NKX2-5* gene mutations and Tibetan nationality has rarely been reported.

Materials and methods

Subjects

Data from a total of 70 Tibetan patients with CHD (CHD-p1) and 70 Tibetan healthy controls (CHD-c1) at Xining Affiliated Hospital of Qinghai University from 2016-2019 were collected retrospectively. All the subjects' birthplaces and parents' residences were > 2,500 meters above sea level. All family members agreed to be in the study and signed the informed consent form. Inclusion criteria for the experimental group: All patients were diagnosed clinically by color Doppler echocardiography and confirmed by cardiac surgery or cardiac catheterization. No diabetes, lung disease, kidney disease, encephalopathy, liver disease, and no recent history of surgery were present. Inclusion criteria for the control group: Healthy Tibetans who were age- and gender-matched with the experimental group and had no previous history of plateau disease or inherited diseases were selected from the physical examination center as the control group. Routine blood tests, cardiac imaging tests, electrocardiogram (ECG) analyses, and other auxiliary examinations were normal. Exclusion criteria for the experimental group: CHD was diagnosed, but there was no objective evidence, such as an echocardiogram. Other types of genetic disorders, such as facial, spinal, limb, and other defects, were present. Exclusion criteria for the control group: Hypertension or other cardiovascular and cerebrovascular diseases. The study protocols were approved by the Ethics Committee of Qinghai University (Xining, China) and were conducted in accordance with the Declaration of Helsinki. All participants in this study signed informed consent forms.

DNA extraction

Peripheral venous blood (5 mL) was drawn from each participant into an ethylenediaminetetraacetic acid dipotassium anticoagulant tubes and centrifuged at 4000 rpm at 4°C for 10 min. The intermediate white blood cell layer was transferred to a new centrifuge tube. Genomic DNA was extracted from the leukocytes using the Gentra Puregene Blood Kit (Qiagen) according to the manufacturer's instructions, and the extracted genomic DNA was diluted to ~200 ng/μL. The integrity of the genomic DNA was observed by 1.5-2.0% agarose gel electrophoresis for 40 min. Genomic DNA was frozen at -20°C until further use.

Exon primer design and inspection

Primers to the exons of *NKX2-5* were designed according to the literature [9] (**Table 1**). Genomic DNA samples were subjected to polymerase chain reaction (PCR) as follows: 25 μL of 2 × Taq PCR Master Mix, 0.5 μL each of upstream and downstream primers (10 μM), 2 μL of DNA template, and ddH_2O to 50 μL. The PCR reaction conditions were as follows: denaturation at 94°C for 5 min; 34 cycles of denaturation at 94°C for 45 s, annealing at 60°C for 45 s, and extension at 72°C for 45 s; extension at 72°C for 5 min. The PCR products were electrophoresed on 1.5-2% agarose gels at 100 V and 100 mA for 40 min. PCR products were stored at -20°C.

Non-coding region inspection

After subsequent sample collection, the number of samples in the experimental group increased to 103 (CHD-p2) and the number of samples in the control group increased to 267 (CHD-c2). DNA extraction was performed as described above. The single nucleotide polymorphisms (SNPs) in the non-coding regions of *NKX2-5* that have been reported in the literature in the past 10 years and that have been

Table 2. Candidate SNPs information of *NKX2-5* gene

Number	SNP ID	Chromosome	Position	Alleles A/B
1	rs703752	5	173232508	C > A, G
2	rs376790353	5	173233982	C > A
3	rs6882776	5	173237160	G > A
4	rs7704485	5	173805538	C > G, T
5	rs2546741	5	173806653	C > T
6	rs2546734	5	173808250	A > G, T
7	rs2561751	5	173805113	G > A
8	rs2546743	5	173806522	T > C
9	rs2546735	5	173807868	A > G
10	rs118026695	5	173235491	T > C

confirmed to affect the occurrence of certain diseases and pathophysiological processes were collated. The mutation sites in *NKX2-5* were identified by the 1000 Genomes Project (http://www.internationalgenome.org/about), and the tag SNP of *NKX2-5* with a minor allele frequency (MAF) of 0.05 was identified by Haploview 4.2. Based on SNP information released by the National Center for Biotechnology Information of the United States, SNPs located in the non-coding region were selected. Finally, a total of 10 SNPs were selected (**Table 2**). PCR amplification primers and single base extension primers were designed for each candidate tag SNP site using Genotyping Tools and MassARRAY Assay Design software (Note S1, Guide S1). Primers were synthesized and sequencing was performed by Beijing BOAO Biotechnology Co., Ltd. (Table S1).

Statistical methods

Experimental data were analyzed using SPSS19.0 statistical software. The Hardy-Weinberg equilibrium (HWE) test was used to evaluate the sample population. A P value > 0.05 indicated that the population met HWE. The genotype and allele frequencies between the groups were tested by the χ^2 test. The odds ratios (OR), 95% confidence intervals (CIs), and P values were presented. The test level was $\alpha = 0.05$. Linkage disequilibrium and haplotype analysis were calculated using the web-based platform athttp://analysis.bio-x.cn.

Results

NKX2-5 exon sequencing

The CHD-p1 group and the CHD-c1 group *NKX2-5* exon sequence alignments showed that synonymous mutations occurred in codon 63 in exon 1 and codon 606 in exon 2 and that there were no sequence differences in the other exons (**Figure 1**).

The Single Nucleotide Polymorphism Database (dbSNP) was consulted and the registration numbers of the two exon mutations in *NKX2-5* were rs2277923 (exon 1) and rs3729753 (exon 2). HWE analysis found that rs3729753 did not conform to the law and was excluded from subsequent statistical analysis (**Table 3**). No significant difference ($P > 0.05$) in the rs2277923 allele frequency between the CHD-p1 and CHD-c1 groups was found after statistical analysis (**Table 4**).

SNPs in non-coding regions of NKX2-5

We could not successfully design primers to detect rs118026695, thus it was not typed. The remaining candidate SNP sites were typed successfully. HWE results for all sites detected in the CHD-p2 and CHD-c2 groups are shown in **Table 5**. One SNP site did not meet HWE, and it was excluded. The genotype and allele frequencies of the rs6882776 and rs2546741 SNP sites in *NKX2-5* differed significantly between the CHD-p2 and CHD-c2 groups (Table S2). Allele A of rs6882776 and allele T of rs2546741 may protect against CHD (**Table 6**).

SNP locus linkage disequilibrium and haplotype analysis in the non-coding region of NKX2-5

According to the position information of all the SNPs listed in **Table 2**, the *NKX2-5* SNPs were sequenced from front to back for the linkage disequilibrium and haplotype analysis. The linkage disequilibrium of SNPs in *NKX2-5* are shown in **Figure 2**. The red modules indicate strong linkage disequilibrium between adjacent SNPs, and the white modules indicate weak or absent linkage disequilibrium. In the CHD-p2 group, SNPs showed significant linkage disequilibrium (**Figure 2A**), whereas linkage disequilibrium of SNPs was rarely observed in the CHD-c2 group (**Figure 2B**). Our analysis indicates that *NKX2-5* haplotypes CAACTCAG and CGGCTCAG are risk factors for CHD, whereas *NKX2-5* haplotypes CAGCCAG and CAGCTCGG are protective factors for CHD (**Table 7**).

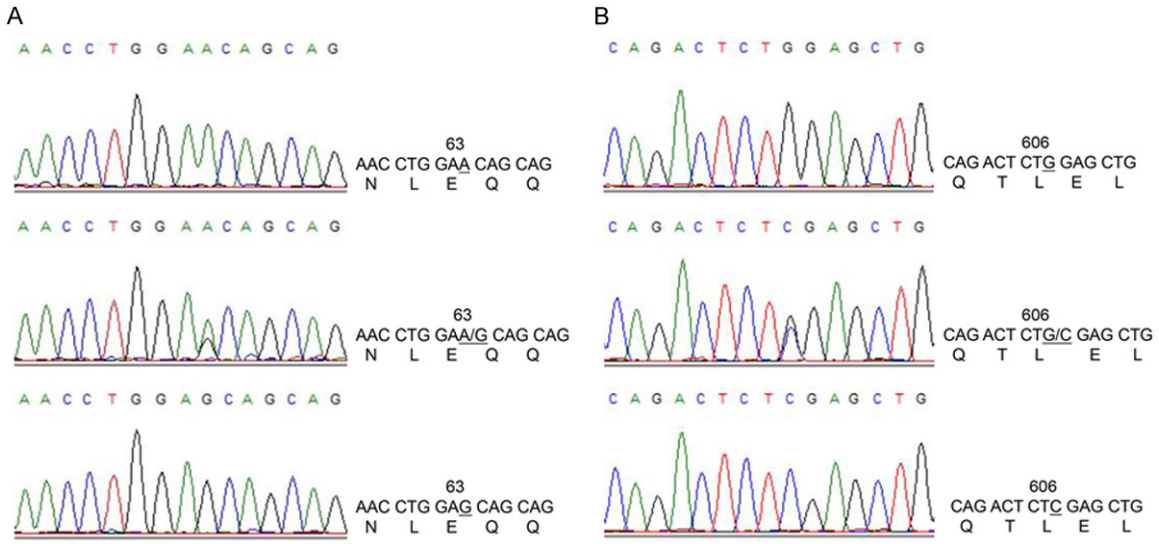

Figure 1. DNA sequencing results. A. Partial nucleotide sequence of exon 1 in the region surrounding codon 63 (rs2277923). B. Partial nucleotide sequence of exon 2 in the region surrounding codon 606 (rs3729753). Notes: N, Asparagine; L, Leucine; E, Glutamic Acid; Q, Glutamine; T, Threonine.

Table 3. HWE analysis of 2 SNPs in the CHD group and control group

SNP ID	*P*	
	CHD-p1	CHD-c1
rs2277923	0.577	0.095
rs3729753	0.855	0.002

Notes: HWE, Hardy-Weinberg Equilibrium; CHD-p1, Tibetan Patients with CHD (group 1); CHD-c1, Tibetan Healthy Controls (group 1).

Discussion

CHD describes congenital malformations with abnormal cardiovascular development in the fetal period, and CHD is the most common birth defect accounting for ~1% of all live births [10]. In addition to suffering from heart disease, patients with CHD have an increased risk of extra-cardiac congenital abnormalities, neurodevelopmental retardation, and a higher cancer rates than the general population [11, 12]. Although the mortality of CHD can be reduced significantly with advanced surgical and medical intervention [13], they require a lot of manpower, material, and financial resources. In addition, although the treatment for CHD has improved, patients remain at risk for cardiac arrhythmias, heart failure, neurodevelopmental deficits, and other congenital anomalies [14]. Further, postoperative sequelae remain a significant patient burden [15]. Therefore, the exploration of CHD, its prevention, and treatment have become urgent medical problems to be solved in China.

Previous studies have shown that environmental factors and genetic factors are risk factors for CHD. Epidemiological surveys have shown that incidences of CHD display significant regional differences [16]. Wu et al. found that the prevalence of CHD in the Qinghai-Tibet plateau was 1.15%, which was significantly higher than the prevalence in the low altitudes of China (0.13-0.35%), and that the prevalence increased with elevation with a prevalence of 1.54% in areas > 4000-5000 meters above sea level [8]. In addition, the incidence of CHD among Tibetan nationals reached 4.86% [17].

The partial pressure of atmospheric oxygen in the Qinghai-Tibet plateau is only 60% of that in the low coastal regions [18]. The decrease in the partial pressure of oxygen reduces gas exchange in the lung leading to a series of pathophysiological changes, which directly affect fetal development. Studies have shown that the mean maternal arterial pressure decreased significantly under hypoxic conditions leading to pathological changes, such as increased uterine vascular tension and decreased uterine blood flow, which reduces fetal birth weight [19]. On the other hand, persistent hypoxia may also cause changes in the

Table 4. Comparison of genotype and allele frequency of rs2277923 between the CHD group and control group

rs2277923	Genotype/Allele	CHD-p1 [n (%)]	CHD-c1 [n (%)]	OR (95% CI)	*P*
Genotype	AA	24 (34.3)	22 (31.4)		
	AG	32 (45.7)	28 (40.0)	0.975 (0.444-2.143)	0.950
	GG	14 (20.0)	20 (28.6)	2.057 (0.847-4.994)	0.109
Allele	A	80 (57.1)	72 (51.4)		
	G	60 (42.9)	68 (48.6)	1.259 (0.786-2.017)	0.337

Notes: CHD-p1, Tibetan Patients with CHD (group 1); CHD-c1, Tibetan Healthy Controls (group 1). Data are presented as the odds ratio (OR) and 95% confidence interval (CI).

Table 5. HWE analysis of 9 SNPs in the CHD group and control group of *NKX2-5* gene

SNP ID	*P*	
	CHD-p2	CHD-c2
rs703752	0.829	0.402
rs376790353	< 0.001	0.926
rs6882776	0.660	0.690
rs2561751	0.112	0.981
rs7704485	0.111	0.550
rs2546743	0.721	0.637
rs2546741	0.231	0.420
rs2546735	0.231	0.730
rs2546734	0.320	0.233

Notes: HWE, Hardy-Weinberg Equilibrium; CHD-p2, Tibetan Patients with CHD (group 2); CHD-c2, Tibetan Healthy Controls (group 2).

genetics of Tibetans. Tibetans have lived in the plateau for a long time and have adapted to the plateau, and research on SNPs has revealed several candidate genes associated with Tibetan high-altitude adaptation, for example, the *EDNRA* gene is associated with Tibetan birth weight [20].

More than 40 genes associated with CHD have been identified, and one research study suggested that > 440 genes may associate with CHD [14]. Schott et al. found that the *NKX2-5* gene associates with CHD in 1998, and since then, research on its function has attracted much attention [21]. The *NKX2-5* gene is located on human chromosome 5q35, contains two exons, and encodes 324 amino acids [22]. The *NKX2-5* protein consists of three highly conserved domains: the tinman domain at the N-terminus, the homeodomain-containing domain, and the NK2-specific domain at the C-terminus [23]. The homeodomain-containing domain has DNA-binding activity and functions in transcriptional activation. The NK2-specific domain is a transcriptional activation regulator and participates in *NKX2-5* gene phosphorylation via the nuclear localization signal. The function of the tinman domain remains unknown [24]. The *NKX2-5* gene is a marker of cardiovascular precursor cells and regulates the development of primary and secondary cardiogenic regions, regulates cell proliferation and differentiation, and regulates the development and function of the cardiac conduction system and cardiomyocytes [25]. *NKX2-5* gene mutations cause various types of CHDs, including atrial septal defect, ventricular septal defect, and tetralogy of Fallot [26].

We sequenced the two exons of the *NKX2-5* gene in the experimental group and control group patients and found mutations in exon 1 and exon 2. There was a synonymous mutation (p. E21) from GAA to GAG in the 63rd base of exon 1 of *NKX2-5* with the SNP sequence number rs2277923. A synonymous mutation (p. L202) from CTG to CTC occurred at base 606 of exon 2 with the SNP sequence number rs3729753. After eliminating rs3729753, which did not conform to HWE, the remaining locus, rs2277923, showed no genotypic or allelic frequency difference between the groups ($P > 0.05$). Therefore, *NKX2-5* is not a susceptible gene for CHD in the Tibetan plateau. Rs2277923 and rs3729753 are two SNPs that are frequently studied when analyzing the relationship between *NKX2-5* and CHD. Whether rs2277923 and rs3729753 associate with CHD remains controversial, perhaps due to the different populations studied (different countries, different ethnic groups, and different regions) or differences in sample sizes. Although the final translated protein products of the synonymous mutations caused by

Table 6. Genotype and allele frequency distribution of *NKX2-5* gene candidate SNPs in CHD group and control group

NKX2-5	Genotype/Allele	CHD-p2 [n (%)]	CHD-c2 [n (%)]	OR (95% CI)	*P*
rs6882776					
Genotype	GG	31 (30.1)	52 (19.5)		
	GA	53 (51.5)	128 (47.9)	0.695 (0.402-1.201)	0.191
	AA	19 (18.4)	87 (32.6)	0.366 (0.188-0.713)	0.003
Allele	G	115 (55.8)	232 (43.4)		
	A	91 (44.2)	302 (56.6)	0.608 (0.440-0.840)	0.003
rs2546741					
Genotype	CC	34 (33.0)	102 (38.2)		
	CT	45 (43.7)	131 (49.1)	0.970 (0.580-1.624)	0.909
	TT	24 (23.3)	34 (12.7)	0.472 (0.246-0.905)	0.023
Allele	C	113 (54.9)	335 (62.7)		
	T	93 (45.1)	199 (37.3)	0.722 (0.521-0.999)	0.049

Notes: CHD-p2, Tibetan Patients with CHD (group 2); CHD-c2, Tibetan Healthy Controls (group 2). Data are presented as the odds ratio (OR) and 95% confidence interval (CI).

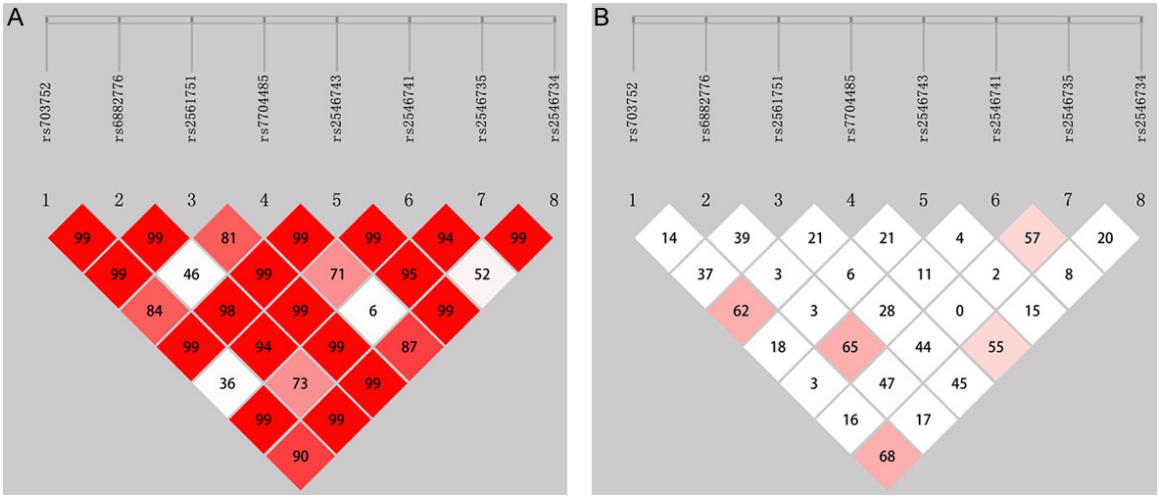

Figure 2. Linkage disequilibrium among the 8 SNPs in *NKX2-5* in (A) the CHD group and (B) the control group.

Table 7. Haplotype analysis for 8 SNPs of the *NKX2-5* gene in the CHD group and control group

Haplotype	CHD-p2 [n (frequency)]	CHD-c2 [n (frequency)]	OR (95% CI)	*P*
AAACTCAA	6.36 (0.031)	10.29 (0.019)	1.664 (0.611-4.537)	0.315
AAATTCAA	22.00 (0.107)	46.99 (0.088)	1.276 (0.746-2.181)	0.373
CAACTCAA	2.00 (0.010)	20.99 (0.039)	0.245 (0.057-1.057)	0.041
CAACTCAG	35.01 (0.170)	60.00 (0.112)	1.674 (1.062-2.638)	0.025
CAGCTCAG	6.50 (0.032)	92.63 (0.173)	0.158 (0.070-0.356)	< 0.001
CAGCTCGG	1.01 (0.005)	56.39 (0.106)	0.043 (0.006-0.306)	< 0.001
CGGCTCAG	31.49 (0.153)	1.93 (0.004)	51.860 (11.972-224.658)	< 0.001
CGGCTTGG	77.13 (0.374)	187.22 (0.351)	1.159 (0.824-1.629)	0.396

Notes: CHD-p2, Tibetan Patients with CHD (group 2); CHD-c2, Tibetan Healthy Controls (group 2). Data are presented as the odds ratio (OR) and 95% confidence interval (CI).

rs2277923 and rs3729753 are the same, the synonymous mutations may not be “silent” as we assumed. For example, Reamon-Buettner et al. found that synonymous mutations nega-

tively affect *NKX2-5* transcriptional activity [7]. In addition, whether the accumulation of several synonymous mutations in the same gene can lead to CHD and whether the severity of CHD can increase or decrease with an accumulation of synonymous mutations remains to be investigated.

Two SNPs in the non-coding region of *NKX2-5*, rs6882776, and rs254674, displayed significant differences between the experimental group and the control group, and all the alleles were protective factors for CHD. After a literature investigation, we found that only a few articles covered these two SNPs. Subsequently, linkage disequilibrium analysis revealed linkage disequilibrium between the 8 candidate tag SNPs. Haplotype analysis showed that three of the eight haplotypes were likely to be protective factors and two were risk factors for CHD. At present, we have conducted many studies on exon sequences and obtained a relatively thorough understanding of them, but we know very little about the non-coding regions. To date, > 90% of disease sites have been found in non-coding regions that do not encode protein [27, 28]. The numbers of SNPs in the non-coding regions of genes are much larger than in the coding regions, and SNPs can affect the functions of transcription factors and cis-acting elements, protein structure, and mRNA transcription leading to diseases. For example, changes in the 5' and 3' untranslated regions can alter the secondary structure of a protein and can alter mRNA regulation, including the amount and location of mRNA in the cell and the rate at which it is translated into protein [29, 30]. In prostate cancer, 88% of mutant SNPs are located in the enhancer [31]. In this study, we found many factors, both protective factors and risk factors, in the non-coding region that may affect the occurrence of CHD. Perhaps "genetic buffering" or "accumulation" can explain this phenomenon. Various positive and negative factors, which may not be obviously related to the development of the heart, may exert an effect when the number of certain factors in the same direction or the degree of action on the heart is greater than that of the opposite factors; the heart would then develop in the direction of this factor. Thus, heredity, environment, and heredity-environment effects should be considered in the pathogenesis of diseases; simple gene analyses may not reveal the cause of the pathogenesis.

Finally, gene analysis plays an important role in medical evaluation because it can detect possible genetic abnormalities associated with CHD and it can allow for individualized drug therapy and clinical diagnosis leading to a good prognosis. Due to the small sample size of the Tibetan population in the plateau, the experimental results may be biased. Sample collection should continue to further improve the accuracy of the experimental results.

Conclusion

SNPs in the exons of *NKX2-5* did not associate with CHD in Tibetans. The rs6882776 and rs2546741 SNPs in the non-coding region of *NKX2-5* may be protective for CHD in the Tibetan population. A haplotype of *NKX2-5* may be associated with CHD in Tibetans.

Acknowledgements

The study is funded by Qinghai Thousand Talents Program, the 2nd Round of "Qinghai 135 High-level Talents Training Project", Basic Applied Study Foundation of Qinghai (2019-ZJ-7006). We also thank all of our colleagues in the Department of Cardiology of the Affiliated Hospital of Qinghai University for their help in sample collection.

Disclosure of conflict of interest

None.

Address correspondence to: Yongnian Liu, Medical College of Qinghai University, 16 Kunlun Road, Xining 810001, Qinghai, China. E-mail: ma9564-17021@163.com

References

[1] Diz OM, Toro R, Cesar S, Gomez O, Sarquella-Brugada G and Campuzano O. Personalized genetic diagnosis of congenital heart defects in newborns. J Pers Med 2021; 11: 562.

[2] Morton SU, Quiat D, Seidman JG and Seidman CE. Genomic frontiers in congenital heart disease. Nat Rev Cardiol 2022; 19: 26-42.

[3] Mullen M, Zhang A, Lui GK, Romfh AW, Rhee JW and Wu JC. Race and genetics in congenital heart disease: application of iPSCs, omics, and machine learning technologies. Front Cardiovasc Med 2021; 8: 635280.

[4] Xu XL, Qi GR, Lu L, Wang Q and Chen QH. A case-control study on influencing factors of congenital heart disease in plateau region of China. Chin Prev Med 2018; 19: 1-4.

[5] Crawford CA, Vujakovich CE, Elmore L, Fleming E, Landis BJ, Spoonamore KG and Ware SM. Adolescents with congenital heart defects: a patient and parental perspective of genetic information and genetic risk. Cardiol Young 2020; 30: 219-226.
[6] Behiry EG, Al-Azzouny MA, Sabry D, Behairy OG and Salem NE. Association of NKX2-5, GATA4, and TBX5 polymorphisms with congenital heart disease in Egyptian children. Mol Genet Genomic Med 2019; 7: e612.
[7] Reamon-Buettner SM, Sattlegger E, Ciribilli Y, Inga A, Wessel A and Borlak J. Transcriptional defect of an inherited NKX2-5 haplotype comprising a SNP, a nonsynonymous and a synonymous mutation, associated with human congenital heart disease. PLoS One 2013; 8: e83295.
[8] Wu TY, Ge RL, Xiao SJ and Wang XZ. Epidemiological study on congenital heart disease in children of Qinghai-Xizang plateau. J Cardiovasc Pulm Dis 1990; 9: 131-134.
[9] Zheng J, Li F, Liu J, Xu Z, Zhang H, Fu Q, Wang J and Sun K. Investigation of somatic NKX2-5 mutations in Chinese children with congenital heart disease. Int J Med Sci 2015; 12: 538-543.
[10] Zaidi S and Brueckner M. Genetics and genomics of congenital heart disease. Circ Res 2017; 120: 923-940.
[11] Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J, Jin SC, Deanfield J, Giardini A, Porter GA Jr, Kim R, Bilguvar K, López-Giráldez F, Tikhonova I, Mane S, Romano-Adesman A, Qi H, Vardarajan B, Ma L, Daly M, Roberts AE, Russell MW, Mital S, Newburger JW, Gaynor JW, Breitbart RE, Iossifov I, Ronemus M, Sanders SJ, Kaltman JR, Seidman JG, Brueckner M, Gelb BD, Goldmuntz E, Lifton RP, Seidman CE and Chung WK. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science 2015; 350: 1262-1266.
[12] Mandalenakis Z, Karazisi C, Skoglund K, Rosengren A, Lappas G, Eriksson P and Dellborg M. Risk of cancer among children and young adults with congenital heart disease compared with healthy controls. JAMA Netw Open 2019; 2: e196762.
[13] Zheng G, Wu J, Chen P, Hu Y, Zhang H, Wang J, Zeng H, Li X, Sun Y, Xu G, Wen S, Cen J, Chen J, Guo Y and Zhuang J. Characteristics of in-hospital mortality of congenital heart disease (CHD) after surgical treatment in children from 2005 to 2017: a single-center experience. BMC Pediatr 2021; 21: 521.
[14] Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, Zeng X, Qi H, Chang W, Sierant MC, Hung WC, Haider S, Zhang J, Knight J, Bjornson RD, Castaldi C, Tikhonoa IR, Bilguvar K, Mane SM, Sanders SJ, Mital S, Russell MW, Gaynor JW, Deanfield J, Giardini A, Porter GA Jr, Srivastava D, Lo CW, Shen Y, Watkins WS, Yandell M, Yost HJ, Tristani-Firouzi M, Newburger JW, Roberts AE, Kim R, Zhao H, Kaltman JR, Goldmuntz E, Chung WK, Seidman JG, Gelb BD, Seidman CE, Lifton RP and Brueckner M. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet 2017; 49: 1593-1601.
[15] McCracken C, Spector LG, Menk JS, Knight JH, Vinocur JM, Thomas AS, Oster ME, St Louis JD, Moller JH and Kochilas L. Mortality following pediatric congenital heart surgery: an analysis of the causes of death derived from the National Death Index. J Am Heart Assoc 2018; 7: e010624.
[16] Li Q, Zhu W, Zhang B, Wu Y, Yan S, Yuan Y, Zhang H, Li J, Sun K, Wang H and Yu T. The MALAT1 gene polymorphism and its relationship with the onset of congenital heart disease in Chinese. Biosci Rep 2018; 38: BSR20171381.
[17] Lv MX, Zheng BH, Yan YF, Li MJ, Zhao LH and Yang J. Screening of congenital heart disease in native Tibetans at high altitude. J Qiqihar Med Univ 2013; 34: 3515-3516.
[18] Wei D, Wei L, Li X, Wang Y and Wei L. Effect of hypoxia on Ldh-c expression in somatic cells of plateau pika. Int J Environ Res Public Health 2016; 13: 773.
[19] Ducsay CA, Goyal R, Pearce WJ, Wilson S, Hu XQ and Zhang L. Gestational hypoxia and developmental plasticity. Physiol Rev 2018; 98: 1241-1334.
[20] Bigham AW, Julian CG, Wilson MJ, Vargas E, Browne VA, Shriver MD and Moore LG. Maternal PRKAA1 and EDNRA genotypes are associated with birth weight, and PRKAA1 with uterine artery diameter and metabolic homeostasis at high altitude. Physiol Genomics 2014; 46: 687-697.
[21] Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, Maron BJ, Seidman CE and Seidman JG. Congenital heart disease caused by mutations in the transcription factor NKX2-5. Science 1998; 281: 108-111.
[22] Khatami M, Mazidi M, Taher S, Heidari MM and Hadadzadeh M. Novel point mutations in the NKX2.5 gene in pediatric patients with non-familial congenital heart disease. Medicina (Kaunas) 2018; 54: 46.
[23] Wang H, Liu Y, Li Y, Wang W, Li L, Meng M, Xie Y, Zhang Y, Yunfeng Z, Han S, Zeng J, Hou Z and Jiang L. Analysis of NKX2-5 in 439 Chinese patients with sporadic atrial septal defect. Med Sci Monit 2019; 25: 2756-2763.

[24] Wang H, Liu Y, Han S, Zi Y, Zhang Y, Kong R, Liu Z, Cai Z, Zhong C, Liu W, Li L and Jiang L. Nkx2-5 regulates the proliferation and migration of H9c2 cells. Med Sci Monit 2020; 26: e925388.
[25] Wang HS, Han S, Liu Z, Li XF, Zhong CB, Li LF, Li YX and Jiang LH. Transcription factor NKX2-5: molecular structure and biological function in regulating cardiovascular precursor cells. Chin J Tissue Eng Res 2021; 25: 108-115.
[26] Kalayinia S, Ghasemi S and Mahdieh N. A comprehensive in silico analysis, distribution and frequency of human Nkx2-5 mutations; a critical gene in congenital heart disease. J Cardiovasc Thorac Res 2019; 11: 287-299.
[27] Boix CA, James BT, Park YP, Meuleman W and Kellis M. Regulatory genomic circuitry of human disease loci by integrative epigenomics. Nature 2021; 590: 300-307.
[28] Zhong W, Liu W, Chen J, Sun Q, Hu M and Li Y. Understanding the function of regulatory DNA interactions in the interpretation of non-coding GWAS variants. Front Cell Dev Biol 2022; 10: 957292.
[29] Khan SM, Faisal AM, Nila TA, Binti NN, Hosen MI and Shekhar HU. A computational in silico approach to predict high-risk coding and non-coding SNPs of human PLCG1 gene. PLoS One 2021; 16: e0260054.
[30] Wright CF, Quaife NM, Ramos-Hernández L, Danecek P, Ferla MP, Samocha KE, Kaplanis J, Gardner EJ, Eberhardt RY, Chao KR, Karczewski KJ, Morales J, Gallone G, Balasubramanian M, Banka S, Gompertz L, Kerr B, Kirby A, Lynch SA, Morton JEV, Pinz H, Sansbury FH, Stewart H, Zuccarelli BD; Genomics England Research Consortium, Cook SA, Taylor JC, Juusola J, Retterer K, Firth HV, Hurles ME, Lara-Pezzi E, Barton PJR and Whiffin N. Non-coding region variants upstream of MEF2C cause severe developmental disorder through three distinct loss-of-function mechanisms. Am J Hum Genet 2021; 108: 1083-1094.
[31] Hazelett DJ, Rhie SK, Gaddis M, Yan C, Lakeland DL, Coetzee SG; Ellipse/GAME-ON consortium; Practical consortium, Henderson BE, Noushmehr H, Cozen W, Kote-Jarai Z, Eeles RA, Easton DF, Haiman CA, Lu W, Farnham PJ and Coetzee GA. Comprehensive functional annotation of 77 prostate cancer risk loci. PLoS Genet 2014; 10: e1004102.

Table S1. *NKX2-5* gene PCR reaction primer and single base extension primer

SNP_ID	2nd-PCRP	1st-PCRP	UEP_SEQ	EXT1_SEQ	EXT2_SEQ
rs6882776	ACGTTGGATGAGGAGAATTTGGGTTGGGAG	ACGTTGGATGAGTCTAGTCTGTGGAGGATG	cTTCCTGGGCCTTGAA	cTTCCTGGGCCTTGAAC	cTTCCTGGGCCTTGAAT
rs2546743	ACGTTGGATGGTCTGTGAGTAGAGATCTGC	ACGTTGGATGCTTTGCACTCCCTCTCCAG	ggATGGAGGAAGCATACC	ggATGGAGGAAGCATACCA	ggATGGAGGAAGCATACCG
rs2546735	ACGTTGGATGTTGGGCTCTCACCAGCTTGA	ACGTTGGATGGAGCTCGCTGTAATCTCTTG	ACCAGCTTGAGAATCC	ACCAGCTTGAGAATCCC	ACCAGCTTGAGAATCCT
rs7704485	ACGTTGGATGTGAATCCAGGAGTCCCATCT	ACGTTGGATGGAAGCCAGGTGAAGAAAAGG	AGTCCCATCTTGACTGT	AGTCCCATCTTGACTGTC	AGTCCCATCTTGACTGTT
rs376790353	ACGTTGGATGGCAGGCCATCCCAGGATCG	ACGTTGGATGAATCGCCCGGGCTCCTGCCTT	CGGCCCCGCTGCAGAAAG	CGGCCCCGCTGCAGAAAGG	CGGCCCCGCTGCAGAAAGT
rs2561751	ACGTTGGATGTTGAGTGTACAGCTCAGTGG	ACGTTGGATGAAGAGTTCTGGAAGTGGGTC	aaAGCTCAGTGGTGTTAAGTGTA	aaAGCTCAGTGGTGTTAAGTGTAC	aaAGCTCAGTGGTGTTAAGTGTAT
rs2546741	ACGTTGGATGAGTCAGTGAGACGTATAGTG	ACGTTGGATGTATCCCTCCACTGATCTACC	ttggATCAGATGGTGATAGACAA	ttggATCAGATGGTGATAGACAAA	ttggATCAGATGGTGATAGACAAG
rs703752	ACGTTGGATGTGACCGATCCCACCTCAACA	ACGTTGGATGACTCAGGGTCATGTTGGGAG	cacaTCAACAGCTCCCTGACTCTCG	cacaTCAACAGCTCCCTGACTCTCGC	cacaTCAACAGCTCCCTGACTCTCGG
rs2546734	ACGTTGGATGGGGACAACTGATGGTTTAGC	ACGTTGGATGGCATGTGGAGTGTGTACTTG	attgGCAGATGAGTTCCCTATGAGCA	attgGCAGATGAGTTCCCTATGAGCAC	attgGCAGATGAGTTCCCTATGAGCAT

Table S2. SNP sites are not statistically significant in the non-coding region of *NKX2-5* gene

SNP	Genotype/Allele	CHD-p [n(%)]	CHD-c [n(%)]	OR (95% CI)	*p*
rs703752					
Genotype	CC	69 (67)	204 (76.4)		
	CA	31 (30.1)	57 (21.3)	1.608 (0.960-2.693)	0.070
	AA	3 (2.9)	6 (2.2)	1.478 (0.360-6.070)	0.585
Allele	C	169 (82.0)	465 (87.1)		
	A	37 (18.0)	69 (12.9)	1.475 (0.954-2.283)	0.079
rs376790353					
Genotype	CC	103 (100.0)	264 (98.9)		
	CA	0 (0)	3 (1.1)	-	0.280
	AA	0 (0)	0 (0)		
Allele	C	206 (100.0)	531 (99.4)		
	A	0 (0)	3 (0.6)	-	0.281
rs2561751					
Genotype	GG	46 (44.7)	136 (50.9)		
	GA	40 (38.8)	109 (40.8)	1.085 (0.663-1.776)	0.746
	AA	17 (16.5)	22 (8.2)	2.285 (1.117-4.674)	0.022
Allele	G	132 (64.1)	381 (71.3)		
	A	74 (35.9)	153 (28.7)	1.396 (0.993-1.963)	0.055
rs7704485					
Genotype	CC	78 (75.7)	207 (77.5)		
	CT	21 (20.4)	55 (20.6)	0.987 (0.560-1.738)	0.964
	TT	4 (3.9)	5 (1.9)	0.471 (0.123-1.799)	0.261
Allele	C	177 (85.9)	469 (87.8)		
	T	29 (14.1)	65 (12.2)	0.846 (0.528-1.354)	0.485
rs2546743					
Genotype	TT	96 (93.2)	252 (94.4)		
	TC	7 (6.8)	15 (5.6)	1.225 (0.485-3.097)	0.668
	CC	0 (0)	0 (0)		
Allele	T	199 (96.6)	519 (97.2)		
	C	7 (3.4%)	15 (2.8)	1.217 (0.489-3.029)	0.672
rs2546735					
Genotype	AA	35 (34.0)	59 (22.1)		
	AG	44 (42.7)	136 (50.9)	1.833 (1.070-3.143)	0.027
	GG	24 (23.3)	72 (27.0)	1.780 (0.954-3.319)	0.068
Allele	A	114 (55.3)	254 (47.6)		
	G	92 (44.7)	280 (52.4)	1.366 (0.989-1.887)	0.058
rs2546734					
Genotype	GG	74 (71.8)	181 (67.8)		
	AG	25 (24.3)	81 (30.3)	0.755 (0.447-1.274)	0.292
	AA	4 (3.9)	5 (1.9)	1.957 (0.511-7.490)	0.319
Allele	G	173 (84.0)	443 (83.0)		
	A	33 (16.0)	91 (17.0)	0.929 (0.601-1.435)	0.739