Research Article
Human and Medical Genetics

A novel *PRRX1* loss-of-function variation contributing to familial atrial fibrillation and congenital patent ductus arteriosus

Zun-Ping Ke[1]*, Gao-Feng Zhang[2,3]*, Yu-Han Guo[2,3], Yu-Min Sun[4], Jun Wang[4], Ning Li[5], Xing-Biao Qiu[5], Ying-Jia Xu[2,3] and Yi-Qing Yang[2,3,6,7] (iD)

[1]*Fudan University, Shanghai Fifth People's Hospital, Department of Geriatrics, Shanghai, China.*
[2]*Fudan University, Shanghai Fifth People's Hospital, Department of Cardiology, Shanghai, China.*
[3]*Fudan University, Shanghai Fifth People's Hospital, Center for Complex Cardiac Arrhythmias of Minhang District, Shanghai, China.*
[4]*Fudan University, Shanghai Jing'an District Central Hospital, Department of Cardiology, Shanghai, China.*
[5]*Shanghai Jiao Tong University, Shanghai Chest Hospital, Department of Cardiology, Shanghai, China.*
[6]*Fudan University, Shanghai Fifth People's Hospital, Cardiovascular Research Laboratory, Shanghai, China.*
[7]*Fudan University, Shanghai Fifth People's Hospital, Central Laboratory, Shanghai, China.*

Abstract

Atrial fibrillation (AF) represents the most common type of sustained cardiac arrhythmia in humans and confers a significantly increased risk for thromboembolic stroke, congestive heart failure and premature death. Aggregating evidence emphasizes the predominant genetic defects underpinning AF and an increasing number of deleterious variations in more than 50 genes have been involved in the pathogenesis of AF. Nevertheless, the genetic basis underlying AF remains incompletely understood. In the current research, by whole-exome sequencing and Sanger sequencing analysis in a family with autosomal-dominant AF and congenital patent ductus arteriosus (PDA), a novel heterozygous variation in the *PRRX1* gene encoding a homeobox transcription factor critical for cardiovascular development, NM_022716.4:c.373G>T;p.(Glu125*), was identified to be in co-segregation with AF and PDA in the whole family. The truncating variation was not detected in 306 unrelated healthy individuals employed as controls. Quantitative biological measurements with a reporter gene analysis system revealed that the Glu125*-mutant *PRRX1* protein failed to transactivate its downstream target genes *SHOX2* and *ISL1*, two genes that have been causally linked to AF. Conclusively, the present study firstly links *PRRX1* loss-of-function variation to AF and PDA, suggesting that AF and PDA share a common abnormal developmental basis in a proportion of cases.

Keywords: Cardiac arrhythmia, congenital heart defect, medical genetics, *PRRX1*, reporter gene analysis.

Received: November 30, 2021; Accepted: February 25, 2022.

Introduction

Atrial fibrillation (AF), characteristic of rapid and disorganized electrical activation and inefficient contraction of the atria, is the most common form of clinical dysrhythmia that affects approximately 1% of the general population globally (Tian *et al.*, 2020; Zhang *et al.*, 2021). Its global prevalence is low in individuals aged <40 years, but increases abruptly beyond the age of 65 years, reaching over 10% in subjects ≥80 years old (Huang *et al.*, 2020). The lifetime for development of AF is ~25% in subjects ≥40 years of age and ~37% in those ≥55 years of age (Lloyd-Jones *et al.*, 2014; Weng *et al.*, 2018). Given that about one third of the total AF population is silent or subclinical asymptomatic, the global prevalence of AF is certainly underestimated (Dilaveris and Kennedy, 2017). AF confers a significantly increased risk for ischemic or hemorrhagic stroke, dementia, venous thromboembolism, acute myocardial infarction, congestive heart failure and

premature death with substantial socioeconomic costs (Kornej *et al.*, 2020). Nevertheless, existing therapeutic regimens for AF are considerably limited in effectiveness and seldom curative, which reflects a poor understanding of the molecular mechanisms underpinning this complex supraventricular arrhythmia (Huang *et al.*, 2020).

Epidemiological investigations have revealed that environmental risk factors predispose to the occurrence and perpetuation of AF, such as advancing age, obesity, obstructive sleep apnea, diabetes mellitus, arterial hypertension, valvular heart diseases, coronary artery disease, dilated cardiomyopathy, hyperthyroidism, heart failure, smoking, alcohol consumption, psychological stress and extreme sports (January *et al.*, 2014; Kornej *et al.*, 2020). However, in ~30% of patients, no well-established cardiovascular pathologies or precipitating factors for AF can be identified, which suggests possible genetic basis underlying AF (Ragab *et al.*, 2020). During the past two decades, multiple epidemiological investigations have demonstrated familial aggregation of individuals with AF and the heritability of AF has been estimated to be as high as 62%, highlighting a strong heritable component responsible for AF (Roselli *et al.*, 2020). By genotyping with

Send correspondence to Yi-Qing Yang. Fudan University, Shanghai Fifth People's Hospital, Cardiovascular Research Laboratory, 801 Heqing Road, Shanghai 200240, China. E-mail: yangyiqing@fudan.edu.cn.

*These two authors contributed equally to the study.

a few hundred polymorphic microsatellite markers scattered throughout the genome and genetic linkage analysis of AF families, Brugada and his partners located the first locus for AF at human chromosome 10q22-q24 (Brugada *et al.*, 1997). Subsequently, similar genetic studies linked more genetic loci to AF, including human chromosome 5p13, 5p15, 6q14-16, 10p11-q21 and 20q12-13 (Ellinor *et al.*, 2003; Oberti *et al.*, 2004; Volders *et al.*, 2007; Darbar *et al.*, 2008). By genetical analysis of a large Chinese family inflicted with AF, Chen and his coworkers mapped a new locus for AF to chromosome 11p15.5 and in this chromosomal region discovered the first AF-causative gene, S140G-mutant *KCNQ1*, which encodes an α subunit of voltage-gated potassium channel (Chen *et al.*, 2003). Functional analysis of the S140G-mutant *KCNQ1* unveiled a gain-of-function impact on the currents of *KCNQ1*/KCNE1 and *KCNQ1*/KCNE2 channels, which significantly shorten the action potential duration of atrial myocytes thereby increasing the vulnerability to AF (Chen *et al.*, 2003). Up to now, in addition to the association of ~140 genetic loci with increased predisposition to AF revealed by genome-wide association studies (Kim *et al.*, 2021), rare variations in over 50 distinct genes have been discovered to contribute to AF, amidst which the majority encode cardiac potassium ion channels, sodium channels, gap junction channels, calcium channels, signaling molecules, structural proteins and transcription factors (Choi *et al.*, 2020; Ghazizadeh *et al.*, 2020; Hansen *et al.*, 2020; Huang *et al.*, 2020; Jiang *et al.*, 2020; Ragab *et al.*, 2020; Roselli *et al.*, 2020; van Ouwerkerk *et al.*, 2020; Wu *et al.*, 2020; Yang *et al.*, 2020; Chalazan *et al.*, 2021; Lazarte *et al.*, 2021; Li *et al.*, 2021a, b; Ziki *et al.*, 2021). Interestingly, multiple variations in or near the *PRRX1* gene, has recently been associated with an enhanced susceptibility to AF in humans (Tucker *et al.*, 2017; Guo *et al.*, 2021; Wu *et al.*, 2021). However, due to pronounced genetical heterogeneity, the genetic determinants underlying AF remain largely elusive. This study was sought to identify a novel genetic variation predisposing to AF.

Material and Methods

Recruitment and clinical evaluation of study participants

For this investigation, a three-generation family affected with AF and congenital patent ductus arteriosus (PDA) was identified, from which 18 available family members were enlisted. A total of 306 unrelated healthy volunteers, who had neither AF nor congenital heart defect (CHD), the most common type of birth defects (Oliveira-Brancati *et al.*, 2020; Virani *et al.*, 2021), were enrolled as control subjects. All study participants experienced a comprehensive clinical assessment, including review of medical histories, physical examination, electrocardiography and echocardiography as well as routine laboratory tests. The healthy control individuals were exactly matched with the cases for gender, ethnicity and age. Clinical diagnosis and classification of AF or CHD were made as previously described (January *et al.*, 2014; Benjamin *et al.*, 2019; Wang *et al.*, 2020; Zhao *et al.*, 2021). This case-control

research was carried out in conformity with the ethical tenets outlined in the Declaration of Helsinki and was approved by the Medical Ethics Committee of Shanghai Chest Hospital (with an approval number of KS1101). Prior to collection of peripheral venous blood samples, informed consent was provided by the study participants or their parents.

Whole-exome sequencing and bioinformatical analysis

Genomic DNA was extracted from the venous blood leucocytes of every test person utilizing a genomic DNA extraction kit (Promega, USA). Whole-exome sequencing (WES) and bioinformatical analysis in five affected family members (I-1, II-1, II-6, III-2 and III-4, Figure 1A) and four unaffected family members (I-2, II-2, II-5 and III-1, Figure 1A) were performed as described previously (Di *et al.*, 2020; Qiao *et al.*, 2020; Linhares *et al.*, 2021; Wang *et al.*, 2021; Xian *et al.*, 2021). In brief, each exome library was constructed using 5 μg of genomic DNA from a study subject, enriched by ligation-mediated polymerase chain reaction (PCR) and captured with the SureSelect Human All Exon V6 Kit (Agilent Technologies, USA). Each exome library was sequenced on the Illumina HiSeq 2000 Genome Analyzer (Illumina, USA) by utilizing the HiSeq Sequencing Kit (Illumina, USA) as per the manufacturer's instructions. Raw image files were processed using the software Pipeline (Illumina, USA) to call bases and the sequences of each subject were generated as a set of reads. By using the Burrows-Wheeler Aligner (BWA) software (Li and Durbin, 2009, 2010), sequencing reads were aligned to the sequences of referential human genome (GRCh37/hg19). Variation calling was performed with the SAMtools (Sequence Alignment/Map Tools, version 0.1.18) software (Li *et al.*, 2009; Li, 2011) and the Genome Analysis Toolkit (GATK, version 4.0.10.1) software (McKenna *et al.*, 2010). The genetic variants that passed the pedigree analysis with any reasonable inheritance pattern of AF and PDA were annotated with the ANNOVAR (annotation of variance, version 20170221) software (Wang *et al.*, 2010). Deleterious variations with a minor allele frequency of <0.001 (in such databases as the Genome Aggregation Database and the Single Nucleotide Polymorphism database) annotated by ANNOVAR were selected as candidate disease-causing variants subject to confirmation by Sanger sequencing analysis in the whole family. The entire coding region and splicing donors/acceptors of the gene harboring a confirmed candidate causative variant were PCR-sequenced in all the available family members and 306 unrelated healthy persons. For an identified rare damaging variation, the Single Nucleotide Polymorphism database (https://www.ncbi.nlm.nih.gov/), the 1000 Genomes Project database (https://www.internationalgenome.org), the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php), and the UK Biobank database (https://www.ukbiobank.ac.uk/) were consulted to check whether it was novel.

Construction of recombinant expression plasmids

The recombinant eukaryotic expression plasmid *PRRX1*-pcDNA3.1 expressing human wild-type *PRRX1* was constructed as previously described (Guo *et al.*, 2021). The Glu125*-mutant

PRRX1-pcDNA3.1 was generated by site-targeted mutagenesis utilizing a complimentary pair of primers (forward primer: 5'-GATGCTTTTGTGCGATAAGACCTTGCCCGCC-3'; reverse primer: 5'-GGCGGGCAAGGTCTTATCGCACAAA AGCATC-3') and a site-directed mutagenesis kit (Stratagene, USA) following the manufacturer's instructions. The Glu125*-

mutant *PRRX1*-pcDNA3.1 underwent selection by *DpnI* (NEB, Hitchin, UK) and was confirmed by sequencing analysis. The *SHOX2*-luciferase (*SHOX2*-luc) and *ISL1*-luciferase (*ISL1*-luc) reporter plasmids, which both express firefly luciferase, were created as described elsewhere (Guo *et al.*, 2021).
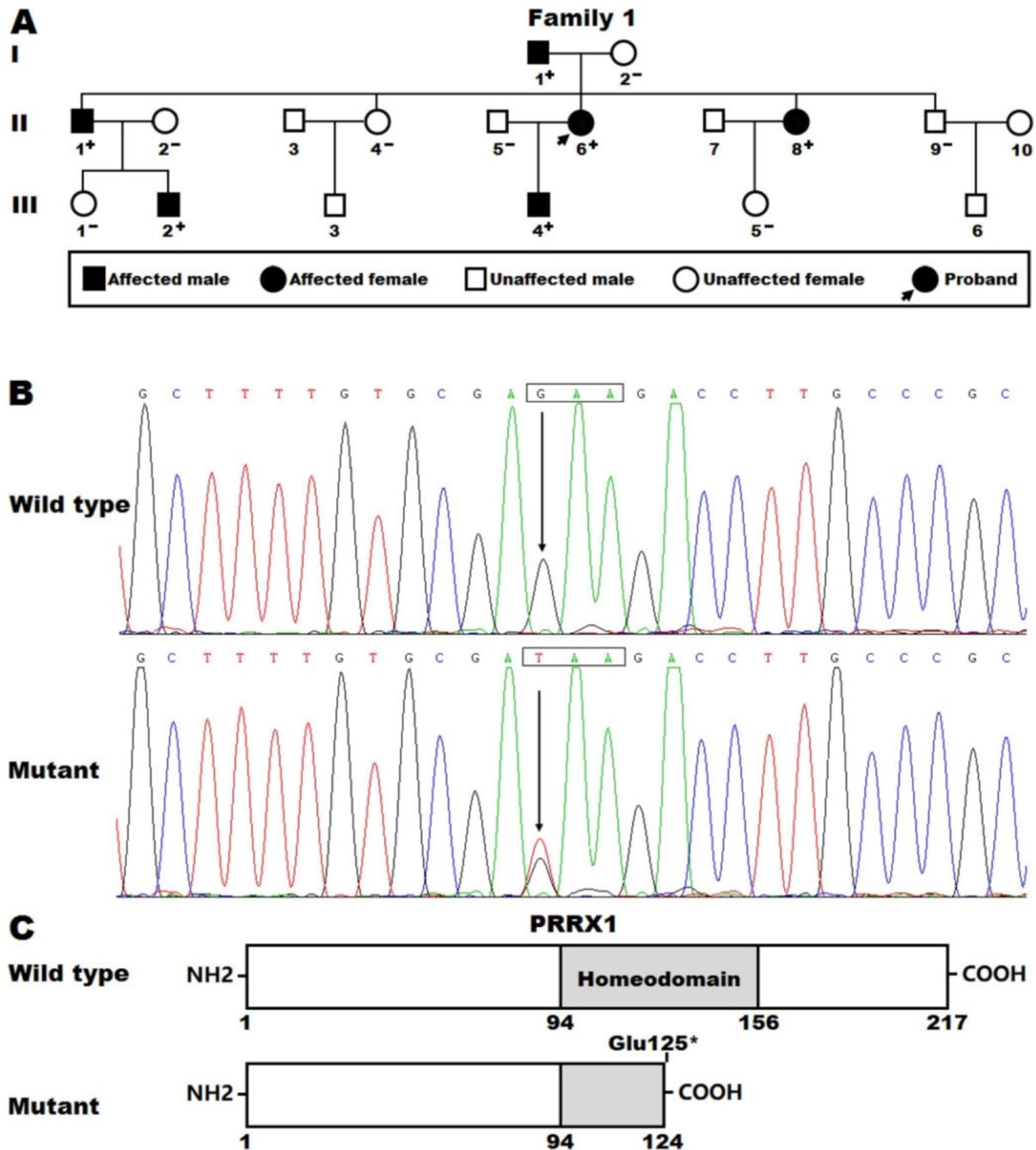

Figure 1 – A new *PRRX1* variation predisposing to familial atrial fibrillation and congenital heart defect. (A) Pedigree structure of the family inflicted with atrial fibrillation and congenital heart disease. "+", carriers of the *PRRX1* variation; "–", non-carriers. (B) Sequence electropherogram traces showing the heterozygous *PRRX1* variation (mutant) as well as its homozygous wild-type control base (wild type). A rectangle delimits a codon comprising three nucleotides. (C) Schemas exhibiting the structural domains of the *PRRX1* proteins. NH2, amino-terminus; COOH, carboxyl-terminus.

Cellular transfection with expression plasmids and dual-luciferase assay

Hela cells were cultivated in Dulbecco's modified Eagle's medium (Invitrogen, USA) containing 10% fetal calf serum (Thermo Fisher Scientific, USA) together with 1% penicillin/streptomycin (Thermo Fisher Scientific, USA) in an incubator with an air of 5% CO_2 at 37 °C. Hela cells were grown in a 12-well plate at an initial density of 1×10^5 cells per well 24 h before transient transfection. Cells were transfected with various amounts of expression plasmids as described previously (Guo *et al.*, 2021). The plasmid pGL4.75 (Promega, USA) expressing renilla luciferase was co-transfected as an internal control to normalize transfection efficiency. The activities of firefly and renilla luciferases were measured on a luminometer (Promega, USA) employing a dual-luciferase reporter assay kit (Promega, USA). The activity of a promoter was expressed as fold activation of firefly luciferase relative to renilla luciferase. For each plasmid, three independent transfections were performed and the resultant data for promoter activity were presented as mean ± standard deviation (SD) of three independent transfection experiments.

Statistical analysis

Unpaired Student's *t*-test was applied to the comparison of promoter activities between two groups. A two-tailed $p<0.05$ was considered to indicate statistical difference.

Results

Clinical characteristics of the pedigree with AF and PDA

As shown in in Figure 1A, a three-generation pedigree with high incidence of AF and PDA was recruited, which comprised 18 living family members, of whom 6 members, including 4 male members and 2 female members with a mean age of 42 years ranging from 19 to 71 years, were diagnosed with AF and PDA in terms of the electrocardiographic and echocardiographic findings. Within this family, AF and PDA were inherited in an autosomal-dominant mode with complete penetrance. Of note, in this pedigree AF began with paroxysmal episodes, but in two family members (members II-1 and II-6) AF became persistent and in one family member (member I-1) AF became permanent over time. No family members had well-established environmental risk factors prone to AF, such as arterial hypertension, valvular heart disease, coronary heart disease, pulmonary heart disease, hyperthyroidism, diabetes mellitus nor obstructive sleep apnea. The proband (member II-6), a forty-three-year-old female member with nineteen years of AF history, was hospitalized due to recurrent syncope and received a successful radiofrequency ablation of AF. The proband's other affected relatives had a history of taking anti-arrhythmic drugs but none of them underwent interventional treatment for AF at the time of enrollment. Additionally, catheter-based closure of PDA was performed in all the affected family members before 6 years of ages except for family member I-1, who underwent closure of PDA at the age of 22. The unaffected family members (six male members and six female members with an average age of 40 years varying from 15 to 68 years) had neither a history of AF nor a history of CHD, with their electrocardiograms and echocardiograms being normal. The clinical features of the pedigree members suffering from AF and PDA are provided in Table 1.

Discovery of a new causative variation in *PRRX1*

WES was conducted in five affected family members (I-1, II-1, II-6, III-2 and III-4, Figure 1A) and four unaffected family members (I-2, II-2, II-5 and III-1, Figure 1A), generating a mean of 22.9 Gb of sequence for each family member, with ~97% mapping to the referential human genome (GRCh37/hg19) as well as ~73% mapping to the target DNA sequences. An average of 18,632 exonic variations (range 17,105-19,148) per family member passed filtering by inheritance model, of which 12 heterozygous nonsense and missense variations passed filtering by ANNOVAR, shared by the five affected family members and predicted to be pathogenic variants, with minor allele frequencies of <0.001 (as summarized in Table 2). Sanger sequencing analysis in the family revealed that only the variant chr1:170,688,998G>T (GRCh37/hg19: NC_000001.10), equivalent to chr1:170,719,857G>T (GRCh38/hg38: NC_000001.11) or NM_022716.4:c.373G>T;p.(Glu125*) in the *PRRX1* gene, was in co-segregation with AF and PDA in the whole family, with complete penetrance. The electropherogram traces exhibiting the heterozygous *PRRX1* variation as well as its homozygous wild-type sequence (used as a control) are exhibited in Figure 1B. The schemas displaying the homeobox domains of wild-type and mutant *PRRX1* proteins are shown in Figure 1C. The truncating variation was neither observed in 306 unrelated healthy individuals, nor found in the Single Nucleotide Polymorphism database, the Single Nucleotide Polymorphism database, the Human Gene Mutation Database, the 1000 Genomes Project database, and the UK Biobank database, indicating it was a novel variation.

No transcriptional activation of *SHOX2* by the Glu125*-mutant *PRRX1* protein

As shown in Figure 2, in cultured Hela cells overexpressing various recombinant expression plasmids, 200 ng of wild-type *PRRX1* plasmid and the same amount (200 ng) of Glu125*-mutant *PRRX1* plasmid transactivated the promoter of *SHOX2* by 36 folds and 1 fold, respectively (wild-type *PRRX1* vs. Glu125*-mutant *PRRX1*: $t = 15.7809$, $p = 0.00009$). When 100 ng of wild-type *PRRX1* plasmid and the same amount (100 ng) of Glu125*-mutant *PRRX1* plasmid were used together, the induced transcriptional activity was 20-fold (wild-type *PRRX1* plasmid plus empty plasmid vs. wild-type *PRRX1* plasmid plus Glu125*-mutant *PRRX1* plasmid: $t = 6.01821$, $p = 0.00384$).

No transcriptional activation of *ISL1* by the Glu125*-mutant *PRRX1* protein

As shown in Figure 3, in cultured Hela cells overexpressing various recombinant expression plasmids, 100 ng of wild-type *PRRX1* plasmid and the same amount (100 ng) of Glu125*-mutant *PRRX1* plasmid transactivated the promoter of *ISL1* by 72 folds and 1 fold, respectively (wild-type *PRRX1* vs. Glu125*-mutant *PRRX1*: $t = 29.5856$, $p = 0.00001$). When 50 ng of wild-type *PRRX1* plasmid and the same amount (50 ng) of Glu125*-mutant *PRRX1* plasmid were used in combination, the induced transcriptional activity was 39-fold (wild-type *PRRX1* plasmid plus empty plasmid vs. wild-type *PRRX1* plasmid plus Glu125*-mutant *PRRX1* plasmid: $t = 11.3859$, $p = 0.00034$).

Table 1 – Clinical features of the family members with atrial fibrillation and patent ductus arteriosus caused by the *PRRX1* variation, NM_022716.4:c.373G>T;p.(Glu125*).

Subject information			Phenotype		Electrocardiogram			Echocardiogram	
Identity (Family 1)	Gender	Age (years)	AF (clinical classification)	CHD (anatomic type)	Heart rate (beats/min)	QRS interval (ms)	QTc (ms)	LAD (mm)	LVEF (%)
I-1	M	71	Permanent	PDA	81	113	464	45	58
II-1	M	48	Persistent	PDA	91	84	432	38	60
II-6	F	43	Persistent	PDA	93	90	445	36	64
II-8	F	40	Paroxysmal	PDA	79	109	416	37	66
III-2	M	21	Paroxysmal	PDA	118	84	468	34	68
III-4	M	19	Paroxysmal	PDA	88	78	427	32	65

AF, atrial fibrillation; CHD, congenital heart disease; F, female; LAD, left atrial diameter; LVEF, left ventricular ejection fraction; M, male; PDA, patent ductus arteriosus; QTc, corrected QT interval.

Table 2 – Nonsynonymous variations in the candidate genes for familial atrial fibrillation and congenital patent ductus arteriosus identified by whole-exome sequencing analysis.

Chr	Position (GRCh37/hg19)	Ref	Alt	Gene	Variant
1	45,101,789	C	G	RNF220	NM_018150.4: c.1081C>G; p.(Gln361Glu)
1	170,688,998	G	T	*PRRX1*	NM_022716.4: c.373G>T; (p.Glu125*)
1	219,383,960	T	A	LYPLAL1	NM_138794.5: c.448T>A; p.(Phe150Ile)
2	125,204,482	G	T	CNTNAP5	NM_130773.4: c.886G>T; p.(Gly296Cys)
3	44,612,251	C	T	ZKSCAN7	NM_018651.4: c.1649C>T; p.(Pro550Leu)
4	140,811,628	G	T	MAML3	NM_018717.5: c.962C>A; p.(Pro321His)
6	155,574,141	G	A	TIAM2	NM_012454.4: c.4179G>A; p.(Trp1393*)
10	13,838,538	A	G	FRMD4A	NM_018027.5: c.257T>C; p.(Phe86Ser)
12	66,742,977	T	G	GRIP1	NM_021150.4: c.3053T>G; p.(Leu1018Arg)
15	29,385,289	T	C	APBA2	NM_005503.3: c.1081T>C; p.(Cys361Arg)
17	4,100,789	C	A	ANKFY1	NM_016376.5: c.982C>A; p.(His328Asn)
19	54,080,349	A	T	ZNF331	NM_018555.6: c.535A>T; p.(Lys179*)

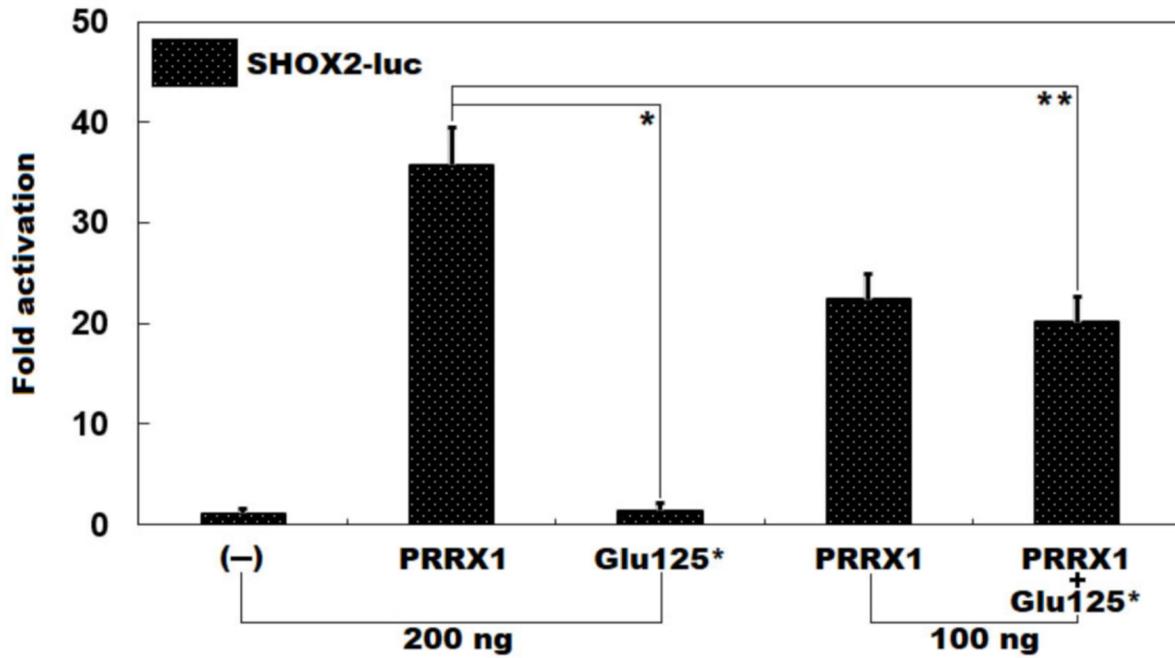Chr, chromosome; Ref, reference; Alt, alteration.

Figure 2 – Failure to transactivate *SHOX2* by Glu125*-mutant *PRRX1*. Dual-luciferase reporter assays unveiled that in cultured Hela cells overexpressing various recombinant expression plasmids, Glu125*-mutant *PRRX1* (Glu125*) failed to transactivate the promoter of the *SHOX2* gene, singly or together with wild-type *PRRX1* (*PRRX1*). The symbols * and ** mean p<0.001 and p<0.005, respectively, in comparison with wild-type *PRRX1* (200 ng).
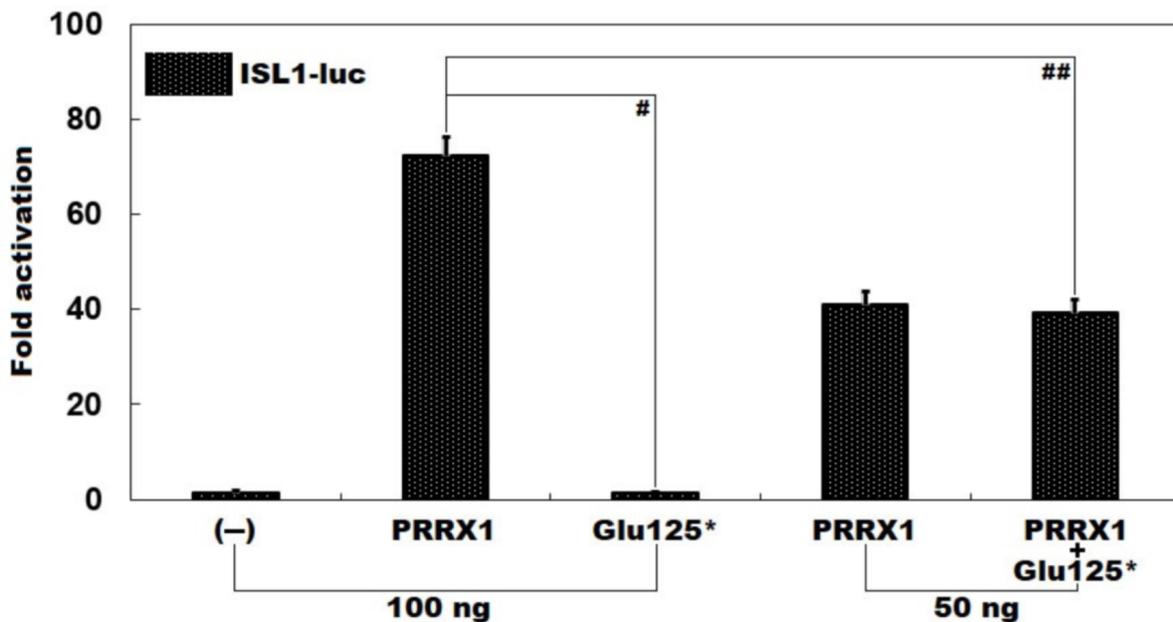

Figure 3 – No transcriptional activation on the promoter of *ISL1* by Glu125*-mutant *PRRX1*. Biological measurement of the transactivation of the *ISL1* promoter-driven luciferase in cultivated Hela cells expressing various expression plasmids by wild-type *PRRX1* (*PRRX1*) or Glu125*-mutant *PRRX1* (Glu125*), alone or in combination, revealed that Glu125* lost the ability to transcriptionally activate the promoter of the *ISL1* gene. Here # and ## mean p<0.0001 and p<0.0005, respectively, in comparison with wild-type counterpart (100 ng).

Discussion

In the present investigation, a heterozygous *PRRX1* variation, NM_022716.4:c.373G>T;p.(Glu125*), was found to be in co-segregation with AF and PDA in a family. The truncating variation was neither detected in 612 referential chromosomes nor retrieved in the Single Nucleotide Polymorphism database, the Single Nucleotide Polymorphism database, the Human Gene Mutation Database, the 1000 Genomes Project database or the UK Biobank database. Functional research unveiled that Glu125*-mutant *PRRX1* lost transcriptional activation on the promoters of *SHOX2* and *ISL1*, two genes where variations have been discovered to result in AF and CHD (Blaschke *et al.*, 2007; Hoffmann *et al.*, 2016, 2019; Li *et al.*, 2018; Ma *et al.*, 2019; Wang *et al.*, 2019; Wu *et al.*, 2020). These findings indicate that genetically defective *PRRX1* contributes to AF and PDA in this family.

In humans, *PRRX1* is localized to chromosome 1q24.2 and encodes paired related homeobox 1, as a member of the paired homeobox-containing family of transcription factors (Grueneberg *et al.*, 1992). The *PRRX1* protein is highly expressed in the cardiovascular system throughout embryonic development, predominantly in the mesenchymal tissues, including the heart, endothoracic great arteries and pulmonary veins (Leussink *et al.*, 1995; Bergwerff *et al.*, 1998; Libório *et al.*, 2011), a common source of ectopic beats triggering AF in the majority of patients (Haïssaguerre *et al.*, 1998). It has been validated that *PRRX1* regulates the epithelial to mesenchymal transition, a hallmark of human cardiovascular morphogenesis (Ocaña *et al.*, 2012). Notably, in *PRRX1*-knockout mice, cardiovascular developmental malformations occurred, encompassing awkward curvature and abnormal positioning of the aortic arch, an aberrant retro-esophageal right subclavian artery as well as a misdirected and elongated ductus arteriosus, highlighting the crucial role of *PRRX1* in the proper development of vessels and perivascular matrices (Bergwerff *et al.*, 2000). Moreover, a recent study has demonstrated that *PRRX1* physically binds to the promoters of *SHOX2* and *ISL1* and transcriptionally activates the expression of *SHOX2* and *ISL1* (Guo *et al.*, 2021), two key downstream target genes responsible for the normal development of the heart, especially for its pacing and conducting system (Blaschke *et al.*, 2007; Liang *et al.*, 2015; Vedantham *et al.*, 2015; Galang *et al.*, 2020) and variations in both *SHOX2* and *ISL1* have been causally linked to AF and CHD (Blaschke *et al.*, 2007; Hoffmann *et al.*, 2016, 2019; Li *et al.*, 2018; Ma *et al.*, 2019; Wang *et al.*, 2019; Wu *et al.*, 2020). In the present research, a new *PRRX1* loss-of-function variation was discovered to lead to AF and PDA. Collectively, these observational results support that *PRRX1* haploinsufficiency is involved in the molecular pathogenesis of AF and CHD in some cases.

Recently, multiple genome-wide association studies and a meta-analysis consistently revealed that a common single nucleotide polymorphism (rs3903239) about 63 kb upstream of the *PRRX1* gene, a top genetic variation at the locus of AF on chromosome 1q24, was associated with significantly increased risk of AF in both Europeans and Asians (Tucker *et al.*, 2017; Wu *et al.*, 2021). Functional analyses unveiled

that this variant diminished the transcriptional activity of the promoter of *PRRX1*, resulting in reduced expression of *PRRX1* in human left atrial tissue (Tucker *et al.*, 2017). Moreover, loss of *PRRX1* was shown to shorten the action potential duration as well as effective refractory period in human atrial cardiomyocytes and zebrafish embryonic myocardium, forming a substrate vulnerable to AF (Tucker *et al.*, 2017), which was further substantiated in a mouse model with deletion of the noncoding AF-associated genomic region (Bosada *et al.*, 2021). Additionally, two loss-of-function variations in *PRRX1* have been uncovered to cause familial AF (Guo *et al.*, 2021). In this research, a new *PRRX1* loss-of-function variation was identified to give rise to AF and PDA, therefore expanding the phenotypic spectrum linked to *PRRX1* and supporting *PRRX1* as a causative gene for AF and CHD. Notably, heterozygous loss-of-function variations in *PRRX1* have already been described in patients with agnathia-otocephaly complex, a rare condition characterized by mandibular hypoplasia or agnathia, ear anomalies (melotia/synotia) and microstomia with aglossia (Dubucs *et al.*, 2021). It is interesting that the same kind of variation has been associated with a quite different phenotype (AF) in the present study, which may be explained in part by the distinct genetic backgrounds.

Conclusions

This study firstly associates *PRRX1* loss-of-function variation with AF and PDA in humans, which suggests that AF and PDA may share a common basis of anomalous cardiovascular development in a subset of cases, implying potential implications for early precise prophylaxis and improved prognostic risk stratification of patients affected with AF and PDA.

Acknowledgements

Conflict of Interest

None.

Author Contributions

ZPK, GFZ, YQY designed the study. ZPK, GFZ, YMS, JW, NL, XBQ, YJX, YQY conducted clinical investigations. ZPK, GFZ, YHG, YMS, YQY completed the biochemical. All authors have approved the submission of the manuscript.

References

ABenjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Chang AR, Cheng S, Das SR *et al.* (2019) Heart disease and stroke statistics–2019 update: A report from the American Heart Association. Circulation 139:e56-e528.

Bergwerff M, Gittenberger-de Groot AC, DeRuiter MC, van Iperen L, Meijlink F and Poelmann RE (1998) Patterns of paired-related homeobox genes PRX1 and PRX2 suggest involvement

in matrix modulation in the developing chick vascular system. Dev Dyn 213:59-70.

Bergwerff M, Gittenberger-de Groot AC, Wisse LJ, DeRuiter MC, Wessels A, Martin JF, Olson EN and Kern MJ (2000) Loss of function of the Prx1 and Prx2 homeobox genes alters architecture of the great elastic arteries and ductus arteriosus. Virchows Arch 436:12-19.

Blaschke RJ, Hahurij ND, Kuijper S, Just S, Wisse LJ, Deissler K, Maxelon T, Anastassiadis K, Spitzer J, Hardt SE *et al.* (2007) Targeted mutation reveals essential functions of the homeodomain transcription factor *SHOX2* in sinoatrial and pacemaking development. Circulation 115:1830-1838.

Bosada FM, Rivaud MR, Uhm J-S, Verheule S, van Duijvenboden K, Verkerk AO, Christoffels VM and Boukens BJ (2021) A variant noncoding region regulates *PRRX1* and predisposes to atrial arrhythmias. Circ Res 129:420-434.

Brugada R, Tapscott T, Czernuszewicz GZ, Marian AJ, Iglesias A, Mont L, Brugada J, Girona J, Domingo A, Bachinski LL *et al.* (1997) Identification of a genetic locus for familial atrial fibrillation. N Engl J Med 336:905-911.

Chalazan B, Mol D, Darbar FA, Ornelas-Loredo A, Al-Azzam B, Chen Y, Tofovic D, Sridhar A, Alzahrani Z, Ellinor P *et al.* (2021) Association of rare genetic variants and early-onset atrial fibrillation in ethnic minority individuals. JAMA Cardiol 6:811-819.

Chen Y-H, Xu S-J, Bendahhou S, Wang X-L, Wang Y, Xu W-Y, Jin H-W, Sun H, Su X-Y, Zhuang Q-N *et al.* (2003) *KCNQ1* gain-of-function mutation in familial atrial fibrillation. Science 299:251-254.

Choi SH, Jurgens SJ, Weng L-C, Pirruccello JP, Roselli C, Chaffin M, Lee CJ-Y, Hall AW, Khera AV, Lunetta KL *et al.* (2020) Monogenic and polygenic contributions to atrial fibrillation risk: Results from a national biobank. Circ Res 126:200-209.

Darbar D, Hardy A, Haines JL and Roden DM (2008) Prolonged signal-averaged P-wave duration as an intermediate phenotype for familial atrial fibrillation. J Am Coll Cardiol 51:1083-1089.

Di R-M, Yang C-X, Zhao C-M, Yuan F, Qiao Q, Gu J-N, Li X-M, Xu Y-J and Yang Y-Q (2020) Identification and functional characterization of KLF5 as a novel disease gene responsible for familial dilated cardiomyopathy. Eur J Med Genet 63:103827.

Dilaveris PE and Kennedy HL (2017) Silent atrial fibrillation: Epidemiology, diagnosis, and clinical impact. Clin Cardiol 40:413-418.

Dubucs C, Chassaing N, Sergi C, Aubert-Mucca M, Attié-Bitach T, Lacombe D, Thauvin-Robinet C, Arpin S, Perez MJ, Cabrol C *et al.* (2021) Re-focusing on Agnathia-Otocephaly complex. Clin Oral Investig 25:1353-1362.

Ellinor PT, Shin JT, Moore RK, Yoerger DM and MacRae CA (2003) Locus for atrial fibrillation maps to chromosome 6q14-16. Circulation 107:2880-2883.

Galang G, Mandla R, Ruan H, Jung C, Sinha T, Stone NR, Wu RS, Mannion BJ, Allu PKR, Chang K *et al.* (2020) ATAC-Seq reveals an *ISL1* enhancer that regulates sinoatrial node development and function. Circ Res 127:1502-1518.

Ghazizadeh Z, Kiviniemi T, Olafsson S, Plotnick D, Beerens ME, Zhang K, Gillon L, Steinbaugh MJ, Barrera V, Sui SH *et al.* (2020) Metastable atrial state underlies the primary genetic substrate for MYL4 mutation-associated atrial fibrillation. Circulation 141:301-312.

Grueneberg DA, Natesan S, Alexandre C and Gilman MZ (1992) Human and Drosophila homeodomain proteins that enhance the DNA-binding activity of serum response factor. Science 257:1089-1095.

Guo X-J, Qiu X-B, Wang J, Guo Y-H, Yang C-X, Li L, Gao R-F, Ke Z-P, Di R-M, Sun Y-M *et al.* (2021) *PRRX1* loss-of-function mutations underlying familial atrial fibrillation. J Am Heart Assoc 10:e023517.

Haïssaguerre M, Jaïs P, Shah DC, Takahashi A, Hocini M, Quiniou G, Garrigue S, Le Mouroux A, Le Métayer P and Clémenty J (1998) Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. N Engl J Med 339:659-666.

Hansen TH, Yan Y, Ahlberg G, Vad OB, Refsgaard L, Dos Santos JL, Mutsaers N, Svendsen JH, Olesen MS, Bentzen BH *et al.* (2020) A novel loss-of-function variant in the chloride ion channel gene CLCN2 associates with atrial fibrillation. Sci Rep 10:1453.

Hoffmann S, Clauss S, Berger IM, Weiß B, Montalbano A, Röth R, Bucher M, Klier I, Wakili R, Seitz H *et al.* (2016) Coding and non-coding variants in the *SHOX2* gene in patients with early-onset atrial fibrillation. Basic Res Cardiol 111:36.

Hoffmann S, Paone C, Sumer SA, Diebold S, Weiss B, Roeth R, Clauss S, Klier I, Kääb S, Schulz A *et al.* (2019) Functional characterization of rare variants in the *SHOX2* gene identified in sinus node dysfunction and atrial fibrillation. Front Genet 10:648.

Huang X, Li Y, Zhang J, Wang X, Li Z and Li G (2020) The molecular genetic basis of atrial fibrillation. Hum Genet 139:1485-1498.

January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC Jr, Conti JB, Ellinor PT, Ezekowitz MD, Field ME *et al.* (2014) 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. Circulation 130:e199-e267.

Jiang W-F, Xu Y-J, Zhao C-M, Wang X-H, Qiu X-B, Liu X, Wu S-H and Yang Y-Q (2020) A novel TBX5 mutation predisposes to familial cardiac septal defects and atrial fibrillation as well as bicuspid aortic valve. Genet Mol Biol 43:e20200142.

Kim JA, Chelu MG and Li N (2021) Genetics of atrial fibrillation. Curr Opin Cardiol 36:281-287.

Kornej J, Börschel CS, Benjamin EJ and Schnabel RB (2020) Epidemiology of atrial fibrillation in the 21st century: Novel methods and new insights. Circ Res 127:4-20.

Lazarte J, Laksman ZW, Wang J, Robinson JF, Dron JS, Leach E, Liew J, McIntyre AD, Skanes AC, Gula LJ *et al.* (2021) Enrichment of loss-of-function and copy number variants in ventricular cardiomyopathy genes in 'lone' atrial fibrillation. Europace 23:844-850.

Leussink B, Brouwer A, el Khattabi M, Poelmann RE, Gittenberger-de Groot AC and Meijlink F (1995) Expression patterns of the paired-related homeobox genes MHox/Prx1 and S8/Prx2 suggest roles in development of the heart and the forebrain. Mech Dev 52:51-64.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987-2993.

Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics 25:1754–1760.

Li H and Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25:2078–2079.

Li N, Wang Z-S, Wang X-H, Xu Y-J, Qiao Q, Li X-M, Di R-M, Guo X-J, Li R-G, Zhang M *et al.* (2018) A *SHOX2* loss-of-function mutation underlying familial atrial fibrillation. Int J Med Sci 15:1564-1572.

Li N, Xu Y-J, Shi H-Y, Yang C-X, Guo Y-H, Li R-G, Qiu X-B, Yang Y-Q and Zhang M (2021a) KLF15 loss-of-function mutation underlying atrial fibrillation as well as ventricular arrhythmias and cardiomyopathy. Genes (Basel) 12:408.

Li R-G, Xu Y-J, Ye WG, Li Y-J, Chen H, Qiu X-B, Yang Y-Q and Bai D (2021b) Connexin45 (GJC1) loss-of-function mutation contributes to familial atrial fibrillation and conduction disease. Heart Rhythm 18:684-693.

Liang X, Zhang Q, Cattaneo P, Zhuang S, Gong X, Spann NJ, Jiang C, Cao X, Zhao X, Zhang X et al. (2015) Transcription factor *ISL1* is essential for pacemaker development and function. J Clin Invest 125:3256-3268.

Libório TN, Acquafreda T, Matizonkas-Antonio LF, Silva-Valenzuela MG, Ferraz AR and Nunes FD (2011) *In situ* hybridization detection of homeobox genes reveals distinct expression patterns in oral squamous cell carcinomas. Histopathology 58:225-233.

Linhares ND, Wilk P, Wątor E, Tostes MA, Weiss MS and Pena SDJ (2021) Structural analysis of new compound heterozygous variants in PEPD gene identified in a patient with prolidase deficiency diagnosed by exome sequencing. Genet Mol Biol 44:e20200393.

Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, D'Agostino RB, Massaro JM, Beiser A, Wolf PA et al. (2014) Lifetime risk for development of atrial fibrillation: The framingham heart study. Circulation 110:1042-1046.

Ma L, Wang J, Li L, Qiao Q, Di R-M, Li X-M, Xu Y-J, Zhang M, Li R-G, Qiu X-B et al. (2019) *ISL1* loss-of-function mutation contributes to congenital heart defects. Heart Vessels 34:658-668.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M et al. (2010) The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297-1303.

Oberti C, Wang L, Li L, Dong J, Rao S, Du W and Wang Q (2004) Genome-wide linkage scan identifies a novel genetic locus on chromosome 5p13 for neonatal atrial fibrillation associated with sudden death and variable cardiomyopathy. Circulation 110:3753-3759.

Ocaña OH, Córcoles R, Fabra A, Moreno-Bueno G, Acloque H, Vega S, Barrallo-Gimeno A, Cano A and Nieto MA (2012) Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer *PRRX1*. Cancer Cell 22:709-724.

Oliveira-Brancati CIF, Ferrarese VCC, Costa AR and Fett-Conte AC (2020) Birth defects in Brazil: Outcomes of a population-based study. Genet Mol Biol 43:e20180186.

Qiao Q, Zhao C-M, Yang CX-, Gu J-N, Guo Y-H, Zhang M, Li R-G, Qiu X-B, Xu Y-J and Yang Y-Q (2020) Detection and functional characterization of a novel MEF2A variation responsible for familial dilated cardiomyopathy. Clin Chem Lab Med 59:955-963.

Ragab AAY, Sitorus GDS, Brundel BBJJM and de Groot NMS (2020) The genetic puzzle of familial atrial fibrillation. Front Cardiovasc Med 7:14.

Roselli C, Rienstra M and Ellinor PT (2020) Genetics of atrial fibrillation in 2020: GWAS, genome sequencing, polygenic risk, and beyond. Circ Res 127:21-33.

Tian X-T, Xu Y-J and Yang Y-Q (2020) Gender differences in arrhythmias: Focused on atrial fibrillation. J Cardiovasc Transl Res 13:85-96.

Tucker NR, Dolmatova EV, Lin H, Cooper RR, Ye J, Hucker WJ, Jameson HS, Parsons VA, Weng L-C, Mills RW et al. (2017) Diminished *PRRX1* expression is associated with increased risk of atrial fibrillation and shortening of the cardiac action potential. Circ Cardiovasc Genet 10:e001902.

van Ouwerkerk AF, Bosada FM, Liu J, Zhang J, van Duijvenboden K, Chaffin M, Tucker NR, Pijnappels D, Ellinor PT, Barnett P et al. (2020) Identification of functional variant enhancers associated with atrial fibrillation. Circ Res 127:229-243.

Vedantham V, Galang G, Evangelista M, Deo RC and Srivastava D (2015) RNA sequencing of mouse sinoatrial node reveals an upstream regulatory role for Islet-1 in cardiac pacemaker cells. Circ Res 116:797-803.

Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Cheng S, Delling FN et al. (2021) Heart disease and stroke statistics-2021 update: A report from the American Heart Association. Circulation 143:e254-e743.

Volders PGA, Zhu Q, Timmermans C, Eurlings PMH, Su X, Arens YH, Li L, Jongbloed RJ, Xia M, Rodriguez L-M et al. (2007) Mapping a novel locus for familial atrial fibrillation on chromosome 10p11-q21. Heart Rhythm 4:469-475.

Wang K, Li M and Hakonarson H (2010) ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38:e164.

Wang S-S, Wang T-M, Qiao X-H, Huang R-T, Xue S, Dong B-B, Xu Y-J, Liu X-Y and Yang Y-Q (2020) KLF13 loss-of-function variation contributes to familial congenital heart defects. Eur Rev Med Pharmacol Sci 24:11273-11285.

Wang T-M, Wang S-S, Xu Y-J, Zhao C-M, Qiao X-H, Yang C-X, Liu X-Y and Yang Y-Q (2021) SOX17 loss-of-function mutation underlying familial pulmonary arterial hypertension. Int Heart J 62:566-574.

Wang Z, Song H-M, Wang F, Zhao C-M, Huang R-T, Xue S, Li R-G, Qiu X-B, Xu Y-J, Liu X-Y et al. (2019) A new *ISL1* loss-of-function mutation predisposes to congenital double outlet right ventricle. Int Heart J 60:1113–1122.

Weng L-C, Preis SR, Hulme OL, Larson MG, Choi SH, Wang B, Trinquart L, McManus DD, Staerk L, Lin H et al. (2018) Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. Circulation 137:10271038.

Wu L, Chu M and Zhuang W (2021) Association between ZFHX3 and *PRRX1* polymorphisms and atrial fibrillation susceptibility from meta-analysis. Int J Hypertens 2021:9423576.

Wu S-H, Wang X-H, Xu Y-J, Gu JN, Yang CX, Qiao Q, Guo XJ, Guo YH, Qiu XB, Jiang WF et al. (2020) *ISL1* loss-of-function variation causes familial atrial fibrillation. Eur J Med Genet 63:104029.

Xian C, Zhu M, Nong T, Li Y, Xie X, Li X, Li J, Li J, Wu J, Shi W et al. (2021) A novel mutation in ext2 caused hereditary multiple exostoses through reducing the synthesis of heparan sulfate. Genet Mol Biol 44:e20200334.

Yang X, Sasano T, Ebana Y, Takeuchi JK, Ihara K, Yamazoe M and Furukawa T (2020) Functional role of the L396R mutation of TKS5 identified by an exome-wide association study in atrial fibrillation. Circ J 84:2148-2157.

Zhang J, Johnsen SP, Guo Y and Lip GYH (2021) Epidemiology of atrial fibrillation: Geographic/ecological risk factors, age, sex, genetics. Card Electrophysiol Clin 13:1-23.

Zhao L, Jiang W-F, Yang C-X, Qiao Q, Xu Y-J, Shi H-Y, Qiu X-B, Wu S-H and Yang Y-Q (2021) SOX17 loss-of-function variation underlying familial congenital heart disease. Eur J Med Genet 64:104211.

Ziki MDA, Bhat N, Neogi A, Driscoll TP, Ugwu N, Liu Y, Smith E, Abboud JM, Chouairi S, Schwartz MA et al. (2021) Epistatic interaction of PDE4DIP and DES mutations in familial atrial fibrillation with slow conduction. Hum Mutat 42:1279-1293.

Internet Resources

The 1000 Genomes Project database, https://www.internationalgenome.org (accessed 20 November 2021).

The Human Gene Mutation Database, http://www.hgmd.cf.ac.uk/ac/index.php (accessed 20 November 2021).

The Single Nucleotide Polymorphism database, https://www.ncbi.nlm.nih.gov/snp (accessed 20 November 2021).

The UK Biobank database, https://www.ukbiobank.ac.uk/ (accessed 20 November 2021).