Somatic mutations in the GATA6 gene underlie sporadic tetralogy of Fallot

RI-TAI HUANG1, SONG XUE1, YING-JIA XU2 and YI-QING YANG3

1Department of Cardiothoracic Surgery, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200127; Departments of 2Cardiology and 3Cardiovascular Research, Shanghai Chest Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200030, P.R. China

Received August 27, 2012; Accepted October 22, 2012

DOI: 10.3892/ijmm.2012.1188

Abstract. Tetralogy of Fallot (TOF) is the most common cyanotic congenital heart disease associated with significant morbidity and mortality in humans. However, the molecular etiology underlying TOF in most patients remains largely unknown. In the present study, sequence analysis of the GATA6 gene was performed from fresh-frozen cardiac tissues and matched blood samples of 52 unrelated patients who underwent surgical repair of TOF. The cardiac tissues and matched blood specimens from 46 patients who underwent cardiac valve replacement due to rheumatic heart disease and blood samples from 200 healthy individuals as controls were genotyped. The functional characteristics of the mutations were assessed using a luciferase reporter assay system. Based on the results, two novel heterozygous GATA6 mutations, p.G367X and p.G394C, were identified in the cardiac tissues of 2 TOF patients, respectively. No mutations were found in the cardiac tissues from 46 patients with rheumatic heart disease and in the blood samples from the 298 participants. Functional analysis demonstrated that the GATA6 mutants were consistently associated with significantly reduced transcriptional activation compared with their wild-type counterpart. This is the first report on the link of somatic GATA6 mutation to TOF, providing novel insight into the molecular mechanism involved in TOF.

Introduction

Congenital heart disease (CHD) is the most common form of birth defect in humans worldwide, affecting nearly 1% of all live births, and is the leading cause of infant mortality from developmental malformations (1). According to specific anatomic or hemodynamic lesions, CHD is clinically classified into at least 21 different types, of which tetralogy of Fallot (TOF), a tetrad of right ventricular outflow tract obstruction, over-riding of the aortic root, ventricular septal defect, and right ventricular hypertrophy, is the most prevalent type of cyanotic CHD, occurring in approximately 3 of every 10,000 neonates alive, and accounts for roughly 7-10% of all congenital cardiac abnormalities. Without surgical repair, 25% of TOF patients with severe obstruction succumb to the disease within the first year, 40% succumb to the disease by the age of 3, 70% by the age of 10, and 95% by the age of 40 (1-3). Various congenital cardiovascular anomalies, such as atrial septal defect, ventricular septal defect, atroventricular septal defect, TOF, patent ductus arteriosus, double outlet right ventricle, aortic stenosis, and transposition of great arteries, can occur alone or together, leading to poor quality of life, cardiac enlargement or hypertrophy, ventricular dysfunction or failure, delayed fetal brain development, pulmonary hypertension, Eisenmenger’s syndrome, arrhythmias, and even sudden cardiac death in the absence of surgical or catheter-based corrections (4-10). Despite the high prevalence and the important clinical significance, the etiology responsible for CHD remains to be identified in an overwhelming majority of patients (11-13).

It is generally understood that normal embryonic heart development is a complex and dynamic process that requires the orchestration of cardiac cell commitment, differentiation, proliferation and migration, and abnormal cardiac morphogenesis appears to be implicated in both environmental and genetic risk factors, which disrupt the finely regulated biological developmental process (11-13). Previous studies demonstrated that an evolutionarily conserved network of transcription factors that connect signaling pathways with genes for muscle growth, patterning, and contractility, including GATA and NK families, plays a pivotal role in early cardiogenesis (14-17), and a great number of germline mutations in NKX2-5, GATA4 and GATA6 have been associated with CHD (18-39). Nevertheless, the genetic determinants underlying CHD remain largely unclear.

Emerging evidence suggests a novel genetic mosaic mechanism for CHD. Somatic mutations have been identified...
in GATA4 and its molecular partners, NKK2-5 and TBX5, as well as the transcription factor HAND1 and HEY2 in cardiac tissue derived from hearts with CHD (40-48). GATA6 is another member of the GATA family, and its expression and function overlap at least partially with those of GATA4, NKK2-5 and TBX5 during cardiovascular genesis, which makes it logical to hypothesize that somatic GATA6 mutations are involved in the pathogenesis of TOF.

In the present study, in order to evaluate the prevalence and spectrum of somatic GATA6 mutations in patients with sporadic TOF and explore the mechanism by which mutated GATA6 predisposes to TOF, the entire coding region and splice junctions of GATA6 was sequenced in patients as comparison to control individuals, and the functional characteristics of the mutant GATA6 were assessed in comparison with its wild-type counterpart by using a luciferase reporter assay system.

Materials and methods

Study subjects. A cohort of 52 unrelated patients with TOF, who underwent cardiac surgery at Shanghai Renji Hospital in China between January 2009 and December 2011, was recruited. Their age ranged from 6 months to 7 years, with an average of 1.24 years at the time of surgery. The patients were evaluated by skillful cardiologists and the diagnosis was made by echocardiography and confirmed by direct view during surgical procedure. The patients with known chromosomal abnormalities or syndromic cardiovascular defects, such as DiGeorge, Alagille, Noonan, Holt-Oram, Marfan and Char syndrome, were excluded from the study. The sample size was adequate to draw the conclusion that the absence of somatic mutations is significant, according to the report by Draus et al (42).

A total of 46 unrelated patients (25 males and 21 females) with rheumatic heart disease undergoing cardiac valve replacement, and 200 unrelated healthy individuals (105 males and 95 females) randomly selected from those undergoing routine physical examinations, were enrolled as controls. In terms of individual medical history and echocardiographic record, the control individuals had no apparent congenital cardiovascular defects, except for subclinical cardiac aberrations such as bicuspid aortic valve or patent foramen ovale.

The participants were Chinese Han people. The ethnic origin of a participant was ascertained by a combination of self-reported ethnicity and a personal questionnaire regarding birthplace, language, religion and ancestry. The study protocol was reviewed and approved by the local Institutional Ethics Committee and written informed consent was obtained from all participants or their guardians prior to the study.

Sample collection and storage. The cardiac muscle tissues from the right ventricular outflow tracts of TOF patients were resected during the routine cardiac surgery procedures. At the time of resection, the discarded cardiac tissue sample was collected after cleaning the blood stain by sterile normal saline and stored at -80°C. Meanwhile, the patient's discarded peripheral venous blood with sodium citrate was collected (the blood samples were mostly used for activated partial thromboplastin time and prothrombin time tests before surgery). The discarded cardiac specimens from the cardiac valves and matched blood samples of the patients undergoing cardiac valve replacement due to rheumatic valvular disease, and the peripheral venous blood samples of healthy individuals were prepared as controls.

DNA extraction. The somatic DNA was extracted from freshly frozen tissues using QIAamp DNA FFPE Tissue kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions. The genomic DNA was extracted from peripheral blood lymphocytes with Wizard Genomic DNA purification kit (Promega, Madison, WI, USA).

Genetic studies. The primers for amplification of the coding exons and flanking splicing sites of the GATA6 gene were designed as previously described (39). Polymerase chain reaction (PCR) was performed in an automated Perkin-Elmer 9700 Thermal Cycler (Applied Biosystems, Foster City, CA, USA). The reaction mixture for PCR used in the present series of experiments consisted of 2 µl of genomic DNA (50-100 ng/µl), 2.5 µl of 10X Taq Buffer, 5 µl of 5X Q Solution, 2 µl of dNTP Mixture (2.5 mM each), 0.5 µl of each primer (20 mM each), 0.25 µl (1.25 U) of HotStar TaqDNA polymerase (Qiagen), and 12.25 µl of deionized H2O, with a final volume of 25 µl. PCR was carried out under the following conditions: pre-denaturation at 95°C for 15 min, followed by 35 cycles of denaturation at 95°C for 1 min, annealing at 62°C for 30 sec, and extension at 72°C for 1 min, and final extension at 72°C for 10 min. To avoid potential carry-over contamination of the PCR mixture, all reactions were performed under stringent conditions as recommended by Kwok and Higuchi (49). Amplified products were analyzed on 1% agarose gels stained with ethidium bromide and purified with QIAquick Gel Extraction kit (Qiagen). Both strands of each PCR product were sequenced with a BigDye® Terminator v3.1 Cycle Sequencing kit under an ABI PRISM 3130XL DNA Analyzer (were from Applied Biosystems). The sequencing primers were the same as previously designed for specific region amplification. The DNA sequences were analyzed with the DNA Sequencing Analysis Software v5.1 (Applied Biosystems). The sequence variation was validated by re-sequencing an independent PCR-generated amplicon from the same subject and met the standard quality control thresholds with a call rate >99%. Additionally, an identified GATA6 sequence variation was searched in the single nucleotide polymorphism (SNP) database of the National Center for Biotechnology Information to confirm the novelty (http://www.ncbi.nlm.nih.gov/SNP).

Alignment of GATA6 protein sequences across species. Multiple GATA6 protein sequences across various species were aligned using the online program of Muscle, version 3.6 (http://www.ncbi.nlm.nih.gov).

Prediction of the causative potential of a GATA6 sequence variation. The pathogenic potential of a GATA6 sequence variation was predicted by MutationTaster (an online program at http://www.mutationtaster.org), which automatically gave a probability for the variation to be either a disease-causing mutation or a benign polymorphism. Notably, the P-value used here is the probability of the correct prediction rather than the probability of error as used in t-test statistics (i.e., a value close to 1 indicates a high ‘security’ of the prediction).
**Plasmids and site-directed mutagenesis.** The recombinant expression plasmid pcDNA3-hGATA6 was kindly provided by Dr Angela Edwards-Ghatnekar, from the Division of Rheumatology and Immunology, Medical University of South Carolina, Charleston, SC, USA. The atrial natriuretic factor (ANF)-luciferase reporter gene, which contains the 2600-bp 5’-flanking region of the ANF gene, i.e., ANF(-2600)-Luc, was kindly provided by Dr Ichiro Shiojima, from the Department of Cardiovascular Science and Medicine, Chiba University Graduate School of Medicine, Chuo-ku, Chiba, Japan. The identified mutation was introduced into the wild-type GATA6 using a QuickChange II XL Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA, USA) with a complementary pair of primers. The mutant was sequenced to confirm the desired mutation and to exclude any other sequence variations.

**Reporter gene assay.** HEK-293 cells were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum. The ANF(-2600)-Luc reporter vector and an internal control reporter plasmid pGL4.75 (hRluc/CMV, Promega) were used in transient transfection assays to evaluate the transcriptional activation function of the GATA6 mutants. HEK-293 cells were transfected with 0.4 µg of wild-type or mutant pcDNA3-hGATA6 expression vector, 0.4 µg of ANF(-2600)-Luc reporter construct, and 0.04 µg of pGL4.75 control reporter vector using Lipofectamine® 2000 Transfection reagent (Invitrogen, Carlsbad, CA, USA). For co-transfection experiments, 0.2 µg of wild-type pcDNA3-hGATA6, 0.2 µg of empty pcDNA3 plasmid or 0.2 µg of mutant pcDNA3-hGATA6, 0.4 µg of ANF(-2600)-Luc, and 0.04 µg of pGL4.75 were used. Firefly luciferase and Renilla luciferase activities were measured with the Dual-Glo luciferase assay system (Promega) 48 h after transfection. The activity of the ANF promoter was presented as fold activation of Firefly luciferase relative to Renilla luciferase. Three independent experiments were performed at minimum for wild-type and mutant GATA6.

**Statistical analysis.** Data are expressed as the means ± SD. Continuous variables were examined for normality of distribution and the Student’s unpaired t-test was used for the comparison of numeric variables between two groups. A comparison of the categorical variables between two groups was conducted using Pearson’s χ² test or Fisher’s exact test when appropriate. A two-tailed P-value <0.05 was considered to indicate a statistically significant difference.

**Results**

**Baseline characteristics of the study population.** A cohort of 52 unrelated Han-nationality patients with sporadic TOF, who underwent cardiac surgery, was enrolled and clinically evaluated as comparison to 46 unrelated Han-race patients with rheumatic heart disease undergoing cardiac valve replacement and 200 ethnically matched, unrelated healthy individuals used as controls. The subjects had neither positive family history of CHD nor established environmental risk factors for CHD, such as maternal illness and drug use in the first trimester of pregnancy, parental smoking, and chronic exposure to toxicants and ionizing radiation. The baseline clinical characteristics of the study population are summarized in Table I.

**Source of samples.** Peripheral venous blood samples were available for all 298 participants. The malformed myocardial tissue samples were obtained from 52 unrelated patients with sporadic TOF who underwent surgical resection of the right ventricular outflow musculature for relieving the stenosis. Generally, the cardiac muscle fragment obtained was ~5x5 mm in size. In addition, the cardiac valvar tissue samples used as
controls were obtained from surgical discards of 46 unrelated patients undergoing cardiac valve replacement.

**GATA6 mutations.** Genomic DNA from the malformed cardiac tissues of 52 TOF patients, the cardiac valvular tissues of 46 patients with rheumatic heart disease, and the peripheral blood lymphocytes of the 298 participants, were screened for GATA6 mutations. As a result, 2 heterozygous mutations in GATA6 were identified in the fresh pathological myocardial tissues of 2 TOF patients, respectively, with a mutational prevalence of ~3.85%. Specifically, a substitution of thymine for guanine in the first nucleotide of codon 367 of the GATA6 gene (c.1099G>T), resulting in a truncated protein with only N-terminal 366 amino acids left (p.G367X), was identified in the cardiac tissue of a 1-year-old male TOF patient. A replacement of guanine by thymine at coding nucleotide 1180 (c.1180G>T), predicting the transition of glycine to cysteine at amino acid position 394 (p.G394C), was identified in the cardiac tissue of a 2-year-old female TOF patient. The sequence chromatograms showing the detected heterozygous GATA6 mutations as comparison to corresponding control sequences are depicted in Fig. 1. A schematic diagram of GATA6 showing the structural domains and the locations of the identified mutations is presented in Fig. 2. The 2 mutations were neither observed in the cardiac valvular tissues of 46 patients with rheumatic heart disease nor found in the peripheral blood samples of the 298 participants. Neither of the 2 mutations was reported in the SNP database at the National Center for Biotechnology Information, which was consulted again on August 20, 2012. Additionally, during a 24-h period of ambulatory electrocardiographic monitoring, no atrial fibrillation was observed in these 2 mutation carriers.

**Alignment of multiple GATA6 protein sequences.** A cross-species alignment of multiple GATA6 protein sequences demonstrated that the affected amino acid p.G394 was completely conserved evolutionarily and the affected amino acid p.G367 was relatively conserved evolutionarily (Fig. 3). However, the amino acids deleted by the p.G367X mutation constitute functionally important structural domains, including 2 zinc fingers and 1 nuclear localization signal (Fig. 2).
Causative potential of GATA6 mutations. The GATA6 mutations of c.1099G>T and c.1180G>T were both automatically predicted to be disease-causing, with P-values of 1.000000 and 0.999997, respectively. No SNPs in the altered regions were found in the MutationTaster database.

Transcriptional activity of the GATA6 mutants. The wild-type GATA6, the G367X-mutant, and the G394C-mutant GATA6 activated the ANF promoter by ~12-, 1- and 3-fold, respectively. When wild-type GATA6 was co-expressed with the same amount of G367X-mutant or G394C-mutant GATA6, the induced activation of the ANF promoter was ~6- or 5-fold, respectively. These results indicate that both GATA6 mutants are associated with significantly reduced activation activity compared with their wild-type counterpart (Fig. 4).

Discussion

Somatic cell is defined as any cell, other than a germ cell, which is involved in the formation of the body, differentiating into the various tissues, organs. Mutations derived from somatic cells can occur in people, but are not transmitted to offspring (50). Diseases associated or causing somatic mutations can be identified by exploring the genetic substance from diseased tissue cells as comparison to that from healthy tissue cells. Somatic mutations can not be detected by genetic analysis of lymphocytic DNA alone, and mosaicism may reduce the likelihood of detection in the affected tissue. Therefore, mutations in cardiac tissues may be absent or sporadic in blood lymphocytes of the same person (40). In the present study, 2 novel heterozygous mutations of GATA6 (p.G367X and p.G394C) were identified in the malformed heart tissues of 2 out of 52 patients with TOF. The mutant alleles were absent in the cardiac tissues of 46 patients with rheumatic heart disease and in the peripheral blood samples of the 298 participants, including 200 ethnically matched healthy individuals. A cross-species alignment of multiple GATA6 protein sequences showed that the altered amino acids were highly conserved evolutionarily. The 2 variants were both predicted to be pathogenic, and the functional experiments substantiated that the mutant GATA6 proteins were associated with significantly decreased or absent transcriptional activity. Therefore, somatic GATA6 mutations may contribute to the pathogenesis of TOF in the 2 mutation-harboring patients.
The current study, no germline GATA6 mutations were discovered. Lin et al (39) made mutational analysis of GATA6 in 21 unrelated patients with persistent truncus arteriosus, and identified 2 heterozygous mutations of p.E486GfsX10 and p.N466H in 2 patients, respectively, with double outlet right ventricle whereas the p.A178V carrier was affected with atrioventricular septal defect, persistent truncus arteriosus, myocardial hypoplasia, reduced myocardial proliferation, and impaired differentiation of vascular smooth muscle cells (53-55). Similarly, compound null of a GATA5 and a GATA6 allele also gives rise to double outlet right ventricle and ventricular septal defect in mice (56). These results from experimental animals demonstrate an exquisite sensitivity of the developing cardiovascular system to the levels of GATA4, GATA5 and GATA6, and indicate that these GATA factors act synergistically to regulate downstream target genes.

Atrial fibrillation has been documented in some CHD patients harboring the germline mutations of GATA4, GATA5 and GATA6 (57-64), which suggests that atrial fibrillation may share a common genetic origin with CHD. However, in the current investigation, no atrial fibrillation was documented in 2 somatic GATA6 mutation carriers, which may be explained by insufficient electrocardiographic monitoring duration of only 24 h for paroxysmal AF, different genetic background, distinct mutational source, delayed or incomplete penetrance, epigenetic modifiers, or environmental factors (61).

In conclusion, this is the first report on the association of somatic GATA6 loss-of-function mutation with increased susceptibility to TOF, which provides novel insight into the molecular pathogenesis of CHD, and suggests the potential implications for the early diagnosis and personalized treatment of CHD.

Acknowledgements

The authors thank the participants for their devotion to the study. This study was supported in part by grants from the National Natural Science Foundation of China (81270161, 81070153 and 30570768), the Personnel Development Foundation of Shanghai, China (2010019), the Natural Science Foundation of Shanghai, China (10ZR1428000), and the Key Program of Basic Research of Shanghai, China (10JC1414002).

References


