

# Rare genetic variants in GATA transcription factors in patients with hypertrophic cardiomyopathy

Cristina Alonso-Montes,<sup>1,2</sup> Julián Rodríguez-Reguero,<sup>2,3</sup> María Martín,<sup>2,3,4</sup> Juan Gómez,<sup>2,5</sup> Eliecer Coto,<sup>2,4,5</sup> Manuel Naves-Díaz,<sup>1,2</sup> César Morís,<sup>2,3,4</sup> Jorge B Cannata-Andía,<sup>1,2,4</sup> Isabel Rodríguez<sup>1,2</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jim-2016-000364>).

<sup>1</sup>Bone and Mineral Research Unit, Instituto Reina Sofía de Investigación, Hospital Universitario Central de Asturias, Oviedo, Spain

<sup>2</sup>Red de Investigación Renal REDinREN from Instituto de Salud Carlos III, Oviedo, Spain

<sup>3</sup>Cardiology Department, Fundación Asturcor, Hospital Universitario Central de Asturias, Oviedo, Spain

<sup>4</sup>Molecular Genetics Unit, Instituto Reina Sofía de Investigación, Hospital Universitario Central de Asturias, Oviedo, Spain

<sup>5</sup>Universidad de Oviedo, Oviedo, Spain

## Correspondence to

Dr Isabel Rodríguez, Bone and Mineral Research Unit, Hospital Universitario Central de Asturias, Edificio FINBA (N+1 F.2), Avda. de Roma s/n, Oviedo 33011, Spain; [irodriguez@hca.es](mailto:irodriguez@hca.es)

Accepted 23 January 2017  
Published Online First  
5 April 2017

Copyright © 2017 American Federation for Medical Research



CrossMark

**To cite:** Alonso-Montes C, Rodríguez-Reguero J, Martín M, et al. *J Invest Med* 2017;**65**:926–934.

## ABSTRACT

Hypertrophic cardiomyopathy (HCM) is a very heterogeneous disease. Although primarily caused by mutations in genes encoding sarcomeric proteins, other genes might explain that heterogeneity. Potential candidate genes are GATA transcription factors that regulate the expression of proteins associated with HCM. Exons of *GATA2*, *GATA4*, and *GATA6* genes were sequenced in 212 patients with unrelated HCM previously analyzed for genes encoding the most frequently mutated sarcomeric proteins. Functional effects of variants were predicted by in silico analyses. 3 potentially pathogenic variants were identified: c.-77G>A in *GATA2*, p.Ala343Thr (rs370588269) in *GATA4*, and p.Pro555Ala (rs146243018) in *GATA6*. Multivariate analyses showed that angina was more frequent in patients carrying sarcomeric and GATA rare variants (55% vs 23.2% in non-carriers of GATA rare variants, OR (95% CI) 7.12 (1.23 to 41.27),  $p=0.029$ ). Among patients without a known causal mutation, GATA rare variants were associated with a greater maximum posterior wall thickness (16.4±4.4 vs 14.0±3.1 mm in non-carriers,  $p=0.021$ ). Thus, variants having a putative effect on GATA genes would alter the expression of their target genes and could modify the hypertrophic response. Therefore, although relatively infrequent in patients with HCM, they may represent a novel insight into the molecular mechanisms related to the pathogenesis of HCM.

## INTRODUCTION

Hypertrophic cardiomyopathy (HCM, MIM #192600) is the most common form of inherited cardiomyopathy and is defined by the presence of increased left ventricular wall thickness, which cannot be explained solely by an abnormal loading condition.<sup>1</sup> The identification of a large number of mutations in genes encoding contractile proteins leads to the consideration of HCM as a ‘disease of the sarcomere’.<sup>2</sup> However, mutations in these genes can explain about 60% of patients with HCM.<sup>1</sup> Importantly, a wide phenotypic heterogeneity is commonly found among carriers of the same mutation, even within a family. Thus, additional genetic variants have been sought in

## Significance of this study

### What is already known about this subject?

- Hypertrophic cardiomyopathy (HCM) is primarily caused by mutations in genes encoding sarcomeric proteins.
- Non-sarcomeric genes have been described as modifiers of HCM features.
- GATA family of transcription factors participates in the hypertrophic response.

### What are the new findings?

- We describe 31 genetic variants, 4 out of them not previously described, in *GATA2*, *GATA4*, and *GATA6* genes in a cohort of 212 patients with hypertrophic cardiomyopathy.
- Three rare variants were identified in patients: c.-77G>A in *GATA2*, p.Ala343Thr in *GATA4*, and p.Pro555Ala in *GATA6*, having putative functional damaging effects in the protein, predicted by in silico analyses.
- The presence of rare variants in GATA, alone or together with sarcomeric mutations, associates with greater maximum posterior wall thickness and angina, respectively.

### How might these results change the focus of research or clinical practice?

- These variants could offer novel insight into the molecular mechanisms related to the pathogenesis of hypertrophic cardiomyopathy (HCM).
- The replication of these findings in bigger populations could provide new genetic variants to use in screening, diagnosis, and prognosis in patients with HCM.
- Understanding the effect of these variants on the expression or function of cardiac proteins could help in the design of new therapeutic drugs.

recent years. Mutations in genes encoding z-disc proteins,<sup>3</sup> calcium handling proteins,<sup>4</sup> or involved in protein degradation<sup>5</sup> have also been related to the pathogenesis of the disease and

can explain an additional relatively small number of HCM cases. Furthermore, polymorphisms in genes of the renin-angiotensin system,<sup>6, 7</sup> in the resistin gene,<sup>8</sup> and in the *MEF2C* gene,<sup>9</sup> among others, have been described as modifiers in HCM features, with controversial results in some cases.<sup>10</sup> These findings support the search for additional genes implicated in the pathophysiology of HCM or in the interindividual variability.<sup>11</sup> Likely candidates are cardiac transcription factors that regulate the expression of cardiac genes involved in the hypertrophic response.<sup>12</sup> Among these regulators is the GATA family, composed of six members in vertebrates. *GATA1*, *GATA2*, and *GATA3* are mainly expressed in hematopoietic cell lineages<sup>13</sup> but also in endocardial and vascular endothelium, where *GATA2* is one of the main regulators of endothelin-1,<sup>14, 15</sup> a recognized prohypertrophic factor.<sup>16</sup> *GATA4*, *GATA5*, and *GATA6* are widely expressed in tissues derived from mesoderm and endoderm, such as heart, gut, liver, and gonads.<sup>17</sup> Interestingly, overexpression of *GATA4* or *GATA6* has been described as necessary and sufficient to induce a hypertrophic response, whereas *GATA4* or *GATA6* knock-down inhibits the development of cardiac hypertrophy.<sup>18, 19</sup> Despite the association between deregulation of GATA transcription factors and cardiac hypertrophy, there are no studies that examine whether their genetic variation is associated with HCM. Thus, the aim of this study was the screening of *GATA2* (MIM #137295), *GATA4* (MIM #600576), and *GATA6* (MIM #601656) genes in a cohort of patients with HCM in order to find genetic variants in these transcription factors that might be associated with the pathophysiology or with the phenotypic heterogeneity of this disorder.

## METHODS

### Patients and controls

A total of 512 unrelated individuals were analyzed. Patients with HCM (n=212) were recruited at the Department of Cardiology at Hospital Universitario Central de Asturias (HUCA), in Oviedo, Spain. All of them were white, from Asturias, a region of the North of Spain with 1 million inhabitants. The HCM diagnosis was performed according to the American College of Cardiology/European Society of Cardiology (ACC/ESC) guidelines.<sup>20</sup> These patients had previously been screened for mutations in the most frequently mutated sarcomeric genes (*MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, *ACTC1*, *TNNC1*, *MYL2*, *MYL3*, and *TPM1*) and their mutations are already published.<sup>21, 22</sup> The main clinical and anthropometric characteristics of the patients are summarized in table 1.

A group of 300 healthy white individuals from Asturias, 76% men, aged 20–75 years (mean±SD: 43.8±12.5 years) were recruited through the Blood Bank and the Department of Cardiology at HUCA and constituted the control population. They were healthy and did not have a history of cardiovascular or systemic diseases but asymptomatic hypertrophy cannot be ruled out.

All samples were obtained in accordance with the declaration of Helsinki guidelines and with ethics approval from the local Ethics Committee for Clinical Investigation of Asturias. All participants signed an informed consent for the study.

**Table 1** Demographic, clinical, and echocardiographic characteristics of the patients with hypertrophic cardiomyopathy included in the study

Men/women (% men)	134/78 (63)
Age (years)	52±17
Age at onset (years)	47±17
BMI (kg/m <sup>2</sup> )	27±5
Family history, n (%)	
Hypertrophic cardiomyopathy	63 (30)
Sudden cardiac death	35 (16.5)
Clinical symptoms, n (%)	
Angina	64 (30)
Syncope	35 (16.5)
Dyspnea	124 (58)
NYHA index	
Class I–II	93 (75)
Class III–IV	31 (25)
Echocardiographic data	
Interventricular septum (mm)	20±5
Posterior wall (mm)	14±3
LVOTO, n (%)	95 (45)
Sarcomeric mutation, n (%)	
MYBPC3	37 (52.9)
MYH7	23 (32.9)
TNNT2	4 (5.7)
TNNI3	2 (2.9)
TPM1	1 (1.4)
ACTC1	1 (1.4)
MYL2	1 (1.4)
MYBPC3 + TNNT2	1 (1.4)

Categorical variables are shown as number of patients (%). Quantitative variables are shown as mean±SD.

BMI, body mass index; LVOTO, left ventricular outflow tract obstruction; NYHA, New York Heart Association functional class.

### Genotyping

Genomic DNA was extracted from peripheral blood samples as previously described.<sup>23</sup> Coding exons, flanking introns, and 5' untranslated region (UTR) of human *GATA2*, *GATA4*, and *GATA6* (GenBank accession numbers NG\_029334, NG\_008177, and NG\_032677, respectively) were amplified using PCR as previously described.<sup>24</sup> Primers were designed to anneal in the flanking introns on the basis of the genomic sequence (primer sequences and specific conditions are available on request). To search for genetic variants, PCR amplicons were then sequenced on both strands using an ABI3130 system, with BigDye chemistry (Applied Biosystems, Foster City, California, USA). Sequences were analyzed with BioEdit Sequence Alignment Editor and compared with reference sequences in public databases from National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>).

### Classification of identified variants

Identified variants were classified as rare if minor allele frequency (MAF) was ≤0.5% in control population. Variants were considered novel if they were not previously reported in public databases (dbSNP; <http://www.ncbi.nlm.nih.gov/>).

snp/; 1000 Genomes Project, <http://www.1000genomes.org/>; or NHLBI Exome Sequencing Project (ESP) database, <http://evs.gs.washington.edu/EVS/>). Variants found in patients were considered as potentially HCM-related variants when they were absent or rare in our control population or a statistically significant difference was found between frequencies in patients and controls, and, in addition, a putative functional effect was described by at least two in silico prediction analysis tools. Cosegregation within the family was also examined for each potentially pathogenic GATA variant.

### In silico sequence analyses

Amino acid sequence conservation among species was analyzed searching the Homologene database (<http://www.ncbi.nlm.gov/HomoloGene>). MutationTaster (<http://www.mutationtaster.org/>) was used to study DNA sequence alterations, analyzing evolutionary conservation, splice-site changes, loss of protein features, and changes that affect the amount of mRNA. The potential pathogenicity of non-synonymous substitutions was predicted using two complementary programs: Sorting Intolerant From Tolerant (SIFT; <http://sift-dna.org>) and Polyphen2 (<http://genetics.bwh.harvard.edu/pph2/>).<sup>25</sup> Identification of possible transcription factor binding motifs affected by the genetic variants located in the 5'UTR was carried out by using the software MatInspector from Genomatix (<http://www.genomatix.de/>) with the default settings (matrix similarity 0.75). Mfold server (<http://mfold.rna.albany.edu/>) was used to evaluate the most thermodynamically stable form of the mRNA, calculating free energies and secondary structures for each variant using the default settings. SNPfold (<http://ribosnitch.bio.unc.edu/snpfold/SNPfold.html>) was used to evaluate the structural consequences of the variation on the mRNA through the computing of a correlation coefficient between the nucleotide pairing probabilities of the two sequences. Both programs were used to analyze changes

located in 5' and 3'UTRs. Non-synonymous variants affecting threonine, serine, or tyrosine residues were studied by NetPhosK 1.0 (<http://www.cbs.dtu.dk/services/NetPhosK/>) to predict potential kinase-specific phosphorylation sites. Server microRNA.org (<http://www.microrna.org/>) and TargetScanHuman 6.0 (<http://www.targetscan.org/>) were used to check whether the 3'UTR variants potentially altered microRNA binding sites.

### Statistical analyses

$\chi^2$  test or Fisher's exact test when necessary was used to compare genotypic and allelic frequencies in patients and controls and deviation from the Hardy-Weinberg equilibrium. Linkage disequilibrium (LD) and haplotype frequencies were calculating with Haploview 4.2.<sup>26</sup> LD blocks were defined based on published criteria.<sup>27</sup> Multivariate linear and logistic regression was used to adjust for demographic characteristics (gender and age). A p value of <0.05 was considered statistically significant. The Bonferroni multiple-testing correction was used when appropriated to adjust p values (0.05 corrected for 22 potential comparisons). Statistical analyses were carried out by Statistical Package for the Social Sciences (SPSS) V17.0 for Windows.

### RESULTS

In the screening of *GATA2*, *GATA4*, and *GATA6* genes, 18 rare (table 2) and 13 common (see online supplementary table S1) genetic variants were identified, with 4 of them never described previously. Nine of the variants were present only in patients with HCM. The genotype distribution of all variants did not deviate from the Hardy-Weinberg equilibrium when tested in controls, except for rs2713604, located in *GATA2*, and therefore, this polymorphism was excluded from further analyses. Individuals with rare variants were all heterozygous.

**Table 2** Rare genetic variants (MAF<0.5% in control population) identified in *GATA2*, *GATA4*, and *GATA6* genes

Gene	Position	dbSNP ID	Nucleotide variant	Amino acid change	Patients	Controls	p Value	ESP	1000G
<i>GATA2</i>	E1	–	c.-77G>A	–	A=0.2	–	–	–	–
		rs116559910	c.*508G>A	–	A=0.2	–	–	–	A=0
<i>GATA4</i>	E2	rs114658939	c.-306C>T	–	T=0.2	–	–	–	T=0
		rs368418329	c.-294G>T	–	T=0.7	T=0.2	0.32	–	–
	E3	rs55788387	c.699G>A	p.Thr233=	A=0.2	–	–	A=0.3	A=0.4
	I3	rs373179740	c.783+51G>A	–	A=0.2	–	–	A=0.2	–
	E4	rs55980825	c.822C>T	p.Cys274=	T=0.2	T=0.2	1.00	T=0.4	T=0.5
	I4	rs147860174	c.909+25G>A	–	A=0.2	A=0.4	1.00	A=0.7	A=0.7
E6	rs370588269	c.1027 G>A	p.Ala343Thr	A=0.2	–	–	A=0.01	–	
	rs3729855	c.1056C>T	p.Asn352=	T=1.4	T=0.2	0.03	T=0.2	T=0	
<i>GATA6</i>	E1	–	c.-181G>A	–	A=0.2	A=0.4	1.00	–	–
	E2	rs758629279	c.296T>C	p.Val99Ala	C=0.2	–	–	–	–
		–	c.849G>A	p.Ala283=	A=0.2	–	–	–	–
	–	rs185325359	c.851C>G	p.Ala284Gly	G=0.2	G=0.2	1.00	–	G=0
	–	–	c.970_978del	p.His324_His326del	del=0.2	–	–	–	–
	I3	rs56314574	c.1302+20C>T	–	T=0.2	T=0.2	1.00	T=0.03	T=0
	E4	rs143026087	c.1374C>T	p.Asn458=	T=0.2	T=0.2	1.00	T=0.3	T=0.1
E7	rs146243018	c.1663C>G	p.Pro555Ala	G=0.5	G=0.2	0.56	G=0.02	G=0.02	

1000G, 1000 Genomes Project in European populations; dbSNP ID: single-nucleotide polymorphism identification in the database dbSNP.

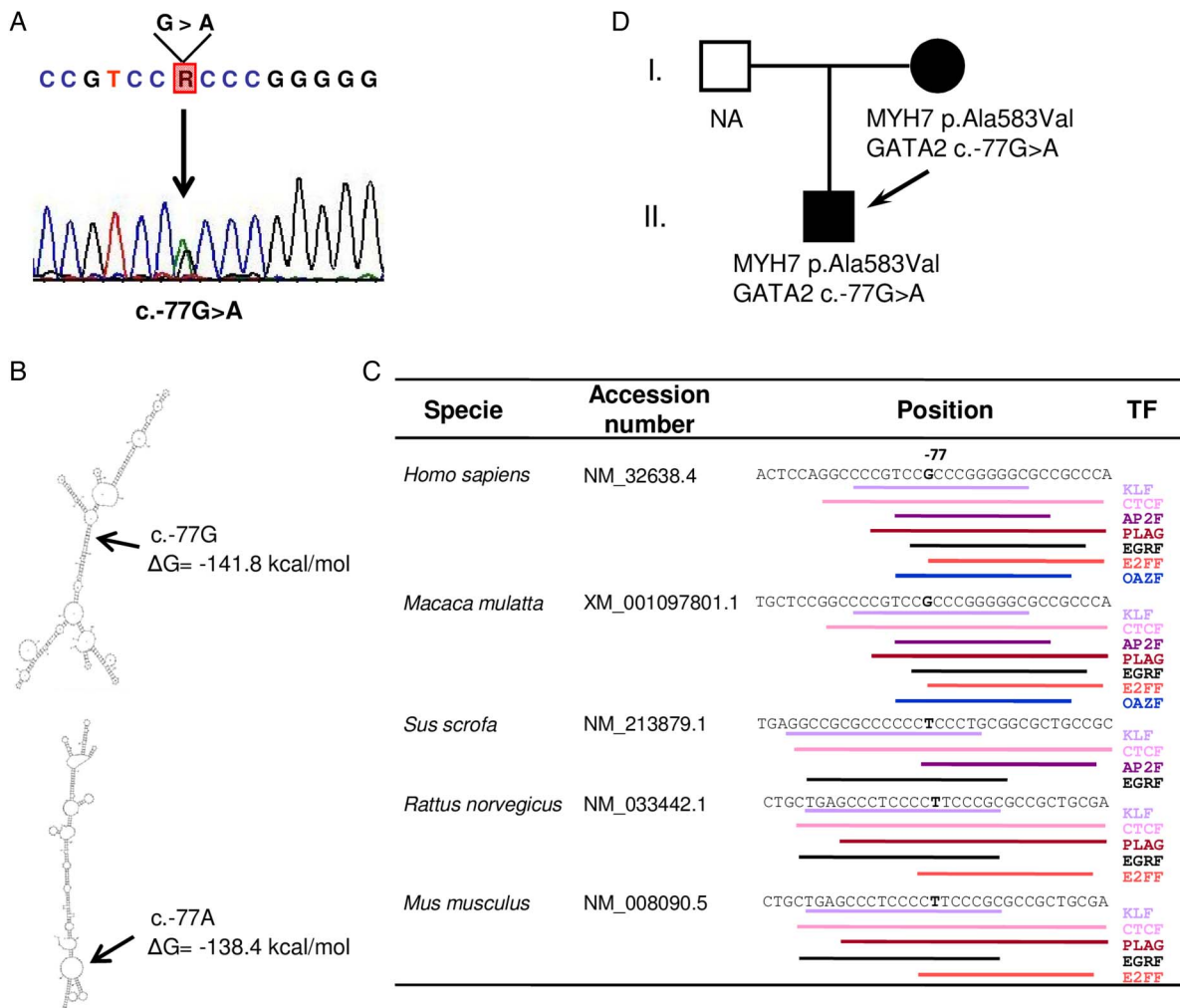
Statistical significance was considered as p value <0.002 (0.05/22). Positions of variants based on *GATA2* RefSeq NM\_032638.4, NP\_116027.2; *GATA4* RefSeq NM\_002052.3, NP\_002043.2; *GATA6* RefSeq NM\_005257.3, NP\_005248.2.

E, exon; ESP, Exome Sequencing Project in European American population; I, intron; MAF, minor allele frequency.

Genetic variants identified in *GATA2*

In the screening of *GATA2*, 11 genetic variants were identified, 1 of them novel (table 2 and see online supplementary table S1). Out of them, 9 were found in patients and healthy subjects and there were no statistical differences in either allele or genotype frequencies between both groups. LD patterns defined a block of 1 kb consisting of rs1108606 and rs3803 polymorphisms, but the estimated haplotype frequencies did not differ between patients and controls (data not shown). Two rare variants were found in two patients (table 2), one variant in each: the novel c.-77G>A in 5'UTR (figure 1A) and c.\*508G>A (rs116559910) in 3'UTR. Analysis of their potential functional effect (see online supplementary table S2) showed that only the variant c.-77G>A is predicted as disease-causing by MutationTaster. A significant alteration on 5'UTR secondary structure was predicted by SNPfold ( $p=0.006$ ; correlation coefficient=0.74) and Mfold

showed that the c.-77A allele changes the secondary structure of the 5'UTR, with only mild alterations in the free energy compared with the c.-77G allele (wild-type allele) (figure 1B). In addition, we also analyzed the putative effect of c.-77G>A in the binding sites for transcription factors with the MatInspector program, which predicted that this variant may create a potential binding site for the Gli-similar Krüppel-like family of transcription factors (GLIs) and abolish recognition motifs for the Activator Protein 2 (AP2) and Krüppel-like factors (KLF). Consensus sequences for AP2 and KLF were located in a region of the promoter conserved across diverse mammals (figure 1C). The patient carrying this variant (a man aged 38 years, carrier of the p.Ala583Val mutation in the sarcomeric gene myosin heavy chain 7 (*MYH7*)) had asymmetric left ventricle hypertrophy with an interventricular septum thickness of 23 mm, who originally presented with syncope and palpitations, developed an NYHA functional class III–IV



**Figure 1** Analyses of the novel variant c.-77G>A in *GATA2*. (A) Electropherogram of c.-77G>A variant (forward sequence). (B) Secondary structure of minimum free energy ( $\Delta G$ ) of c.-77G (wild type) and c.-77A alleles in the 5' UTR of *GATA2* mRNA predicted by Mfold. Arrows indicate the site where the variant c.-77G>A was located. (C) Conservation of putative transcription factors binding sites in the region of c.-77G>A in *GATA2* gene from several species. AP2F, activator protein 2; CTCF, CCCTC-binding factor; E2FF, E2F-myc activator; EGRF, early growth response factors; KLF, Krüppel-like factor; OAZF, Olf1/EBF-associated zinc finger protein; PLAG, pleomorphic adenoma gene. (D) Pedigree of the family carrying c.-77G>A. MYH7: cardiac  $\beta$ -myosin heavy chain; NA: not analyzed; squares, males; circles, females; filled symbols, cases with hypertrophic cardiomyopathy; empty symbols, normal phenotype. The arrow indicates the index patient.

dyspnea, and had episodes of angina. The proband's mother, who also carried both variants (figure 1D), was diagnosed with concentric left ventricle hypertrophy (21 mm of posterior wall thickness) at the age of 50, and presented with atrial fibrillation.

### Genetic variants identified in *GATA4*

The screening of *GATA4* in our cohort revealed 11 variants (table 2 and see online supplementary table S1), 7 of them found in patients and controls, and the analysis of the allele and genotype frequencies showed a statistically significant difference between both groups only for p.Asn352=(rs3729855), a synonymous polymorphism in exon 6. The frequency of T allele and CT genotype were significantly higher in patients with HCM when compared with controls ( $p=0.03$ ) (table 2), an association that persisted after adjusting for sex and age ( $p=0.03$ ; OR=10.78; 95% CI 1.21 to 96.28). However, this association did not survive the Bonferroni correction for multiple comparisons. The analysis of the LD between common variants identified in *GATA4* indicated that they did not constitute haplotype blocks according to defined criteria.<sup>27</sup>

In addition, four rare variants in *GATA4* were found only in patients (table 2). In silico analyses predict a functional effect only for p.Ala343Thr variant (see online supplementary table S2), located in exon 6 of *GATA4* gene, in a residue conserved across mammalian species (figure 2). Thus, SNPfold predicted a potential effect in the structure of mRNA ( $p=0.05$ ; correlation coefficient=0.09). In addition, Netphosk predicted a change in protein phosphorylation pattern identifying a consensus site for protein kinases B and C that could be affected by this change. This variant was found in a sporadic, man aged 56 years, without a sarcomeric mutation identified, and with an interventricular septum thickness of 24 mm. This patient presented with atrial fibrillation detected by ECG and a dilation of the left atrium of 46 mm.

### Genetic variants identified in *GATA6*

Analysis of all seven exons and the flanking introns of *GATA6* in our cohort revealed nine variants, including three not previously described (table 2 and see online

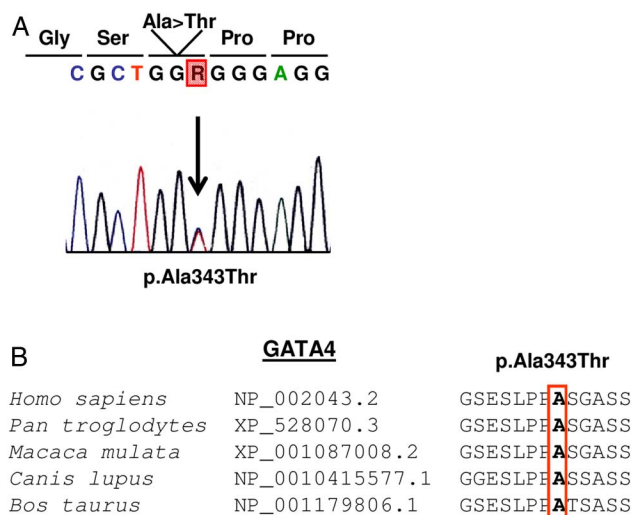
supplementary table S1). Among the variants, eight were considered rare; therefore, LD and haplotype frequencies were not determined for this gene. Allele and genotype frequencies did not differ significantly between patients and controls. However, among the rare variants, the p.Pro555Ala change (figure 3A) was predicted to be pathogenic by MutationTaster and Polyphen-2 tools (see online supplementary table S2). In addition, SNPfold showed a marginally significant p value ( $p=0.06$ ; correlation coefficient=0.90). The altered amino acid was conserved from *Xenopus* to humans (figure 3B). This change was found in two men carrying a sarcomeric mutation and with a family history of HCM and sudden cardiac death, and also in one control. One of these patients is an asymptomatic carrier of the p.Gly236del mutation in the sarcomeric gene cardiac myosin binding protein C (*MYBPC3*), with an asymmetric left ventricle hypertrophy (17 mm of interventricular septum thickness). The other patient who was diagnosed at the age of 51 carried the p.Arg663His mutation in *MYH7* and presented with an asymmetric left ventricle hypertrophy (29 mm of interventricular septum thickness), with left ventricle outflow tract obstruction of 80 mm Hg. This patient developed an NYHA functional class I-II dyspnea, angina, and aortic murmur. In this patient, cosegregation within the family (figure 3C) showed that his sister was diagnosed at the age of 48 with left ventricle hypertrophy, with an interventricular septum thickness of 22 mm and showed fatigue, and also carries both variants. The proband's son also carries both variants, but at the age of 28, he had normal echocardiographic parameters.

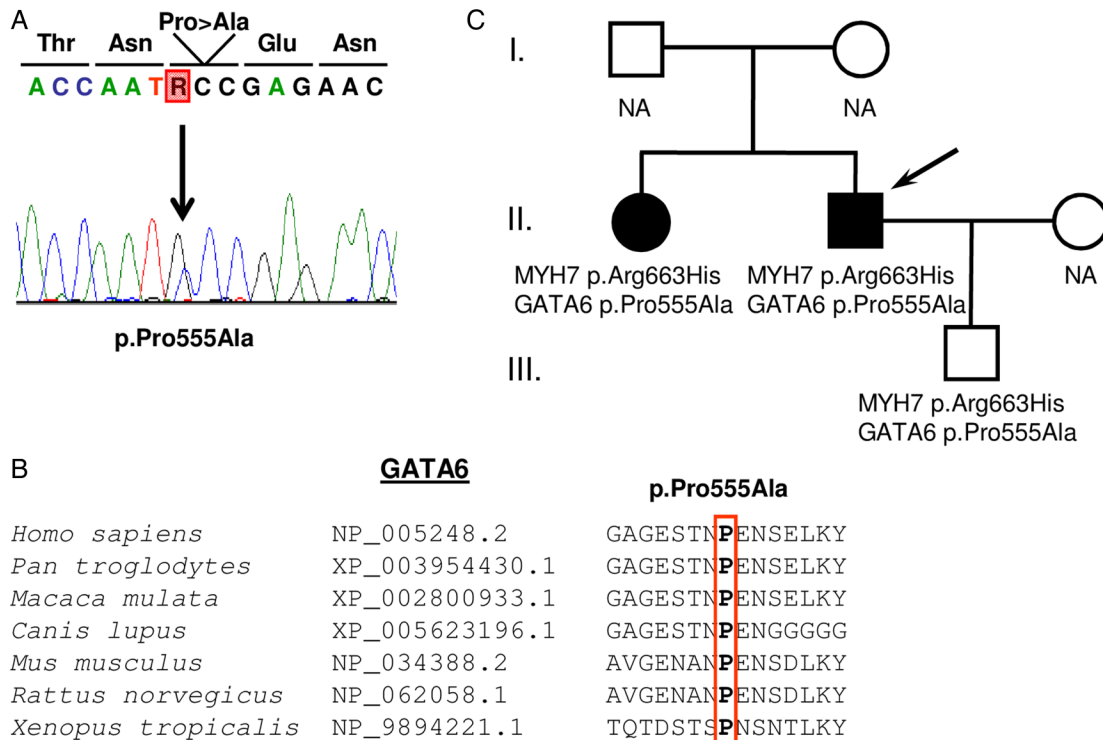
Three rare variants in *GATA6* were found only in patients and two of them involve changes in the amino acid sequence (p.Val99Ala, p.His324\_His326del), although any putative functional effect was determined by in silico analysis.

### Effect of rare variants in *GATA2*, *GATA4*, and *GATA6* on HCM phenotype

In order to determine whether the presence of rare variants in these genes could modify the disease phenotype, the clinical characteristics of patients with HCM carrying the 18 *GATA* rare variants (MAF $\leq$ 0.5% in control population) were analyzed in both carriers of sarcomeric mutation and patients without a known causal mutation. Among the 70

**Figure 2** Analyses of p.Ala343Thr variant in *GATA4*. (A) Electropherogram showing the variant in *GATA4* gene (reverse sequence). (B) Alignment of multiple *GATA4* proteins across mammalian species in the region of p.Ala343Thr variant.





**Figure 3** Analyses of p.Pro555Ala variant in *GATA6*. (A) Electropherogram showing the variant in *GATA6* gene (forward sequence). (B) Alignment of multiple *GATA6* proteins across vertebrate species in the region of p.Pro555Ala variant. (C) Pedigree of the family carrying p.Pro555Ala. MYH7: cardiac  $\beta$ -myosin heavy chain; NA, not analyzed; squares, males; circles, females; filled symbols, cases with hypertrophic cardiomyopathy; empty symbols, normal phenotype. The arrow indicates the index patient.

patients with mutations in sarcomeric genes, 9 carried rare variants in at least 1 of the GATA genes examined (table 3). Multivariate analyses showed that, among the double carriers, angina was more frequent (55.5% vs 23.2%, OR (95% CI) 7.12 (1.23 to 41.27),  $p=0.029$ ) (table 4). In the analysis of patients without sarcomeric mutation, patients carrying rare variants in GATA genes had greater maximum posterior wall thickness after adjusting for sex and age ( $16\pm 4$  vs  $14\pm 3$  mm in non-carriers, standardized coefficient  $\beta=2.39$ ,  $p=0.019$ ) (table 4).

## DISCUSSION

HCM is a genetic cardiac disorder with a high risk for sudden death in undiagnosed patients, which highlights the importance of an accurate diagnosis. Cardiac transcription factors play a key role since they directly regulate an important number of genes whose expression is altered in cardiac hypertrophy.<sup>12</sup> In fact, exclusive overexpression of several of these transcription factors, mainly *GATA4* and *GATA6*,<sup>18 19</sup> but not only,<sup>28 29</sup> is sufficient to induce cardiac hypertrophy, highlighting their importance in the pathogenesis of HCM, which some authors considered, at least in part, as a transcriptional disorder.<sup>30</sup> Therefore, any functional genetic variant affecting these genes could be affecting the developing of HCM. To the best of our knowledge, this is the first analysis of the genetic variation of three members of the GATA family of transcription factors in HCM, although loss-of-function mutations in *GATA4* and *GATA6* have recently been linked to dilated cardiomyopathy.<sup>31 32</sup>

Herein, the screening of *GATA2*, *GATA4*, and *GATA6* in our cohort identified 31 genetic variants, 18 of them with a very low frequency in our population but also in general databases and, therefore, could be considered as pathological candidates.<sup>33</sup> Out of them, three had a predicted functional effect using in silico analyses: one in the 5'UTR of *GATA2* (c.-77G>A) and two non-synonymous changes, one in *GATA4* (p.Ala343Thr), and another in *GATA6* (p.Pro555Ala).

The c.-77G>A variant in *GATA2* was found in a carrier of a sarcomeric mutation and both changes cosegregated with HCM in the family. Then, it could act as phenotype modifier, taking into account that the disease was caused by the sarcomeric mutation. Bioinformatic analyses showed that this variant could affect gene transcription due to the removal of recognition motifs for AP2 and KLF transcription factors. In fact, among KLF family members, KLF15 is an antihypertrophic factor, through the inhibition of *GATA4* and *MEF2* function.<sup>34</sup> In addition, this variant could modify translation efficiency as it could affect the mRNA secondary structure.<sup>35</sup> Furthermore, variants in gene regulatory regions have been previously described as potential modifiers of the HCM phenotype in carriers of sarcomeric mutations, as it was found in *MEF2C*.<sup>9</sup>

Regarding non-synonymous changes, p.Ala343Thr variant was predicted to affect mRNA structure as well as *GATA4* phosphorylation, which is essential to activate the hypertrophic response by increasing its transactivating activity.<sup>36</sup> Furthermore, this change was located in the carboxy-terminal transactivation domain through which

**Table 3** Clinical and echocardiographic characteristics of patients with hypertrophic cardiomyopathy carrying sarcomeric and rare variants (MAF ≤ 0.5% in control population) in GATA genes

GATA gene	Number of carriers	Position	Nucleotide variant	Amino acid change	Sarcomeric gene	Number of carriers	dbSNP ID	Position	Amino acid change	Sex/age	IVS	Clinical profile
2	1	E1	c.-77G>A		MYH7	1	-	E16	p.Ala583Val	M/38	23	Dyspnea (class III-IV), angina, syncope, palpitations
	1	E6	c.*508G>A		MYH7	1	rs767300277	E39	p.Ile1927Phe	F/52	16	Angina
4	1	E2	c.-306C>T		MYH7	1	rs201307101	E32	p.Lys1459Asn	F/56	19	Dyspnea (class III-IV), angina, arrhythmia, murmur, palpitations
	1	E3	c.699G>A	p.Thr233=	MYBPC3	1	rs121909374	E18	p.Glu542Gln	M/57	20	Angina, murmur
	1	E6	c.1056C>T	p.Asn352=	TNNT2	1	rs397516470	E11	p.Glu163del	F/35	28	Dyspnea (class I-II), palpitations
6	1	E1	c.-181G>A		TNNT2	1	rs121964856	E9	p.Arg92Gln	F/-	17	Palpitations
	1	E2	c.851C>G		ACTC1	1	rs193922680	E3	p.Glu101Lys	F/25	17	Asymptomatic
	2	E7	c.1663C>G	p.Ala284Gly	MYH7	1	rs371898076	E18	p.Arg663His	F/51	29	Dyspnea (class I-II), angina, murmur
					MYBPC3	1	-	E7	p.Gly263del	M/49	17	Asymptomatic

dbSNP ID, single-nucleotide polymorphism identification in the database dbSNP; E, exon; F, female; IVS, interventricular septum thickness (mm); M, male. Class refers to New York Heart Association functional class. Positions of variants based on GATA2 RefSeq NM\_032638.4, NP\_116027.2; GATA4 RefSeq NM\_002052.3, NP\_002043.2; GATA6 RefSeq NM\_005257.3, NP\_005248.2; MYH7 RefSeq NP\_000248.2; MYBPC3 RefSeq NP\_000247.2; TNNT2 RefSeq NP\_001263276.1; ACTC1 RefSeq NP\_005150.1.

GATA4 interacts with other transcription factors such as GATA6 and MEF2 to synergistically activate downstream target promoters.<sup>37 38</sup> In addition, the carrier of p. Ala343Thr variation also had atrial fibrillation, which was reported to be linked to GATA4 mutations.<sup>39</sup>

This study also identifies a higher frequency of patients carrying the T allele for p.Asn352=, located in exon 6 of GATA4. This polymorphism was first described in the Glasgow Heart Scan Study, in 95 subjects whose age-adjusted left ventricle mass was in the upper decile of the sex-specific distribution.<sup>40</sup> Although this polymorphism did not survive the Bonferroni correction, the rare allele was over-represented in patients compared with our control subjects and with other populations previously described; therefore, this association should be confirmed in larger populations and then, it could be taken into account in the analyses of other patients with HCM.

The other non-synonymous change, p.Pro555Ala located in GATA6, could have a damaging functional effect as predicted by in silico studies. This change could affect GATA6 protein stability because it is located in the carboxy-terminal domain identified as critical for GATA6 degradation via ubiquitination and proteosomal degradation.<sup>41</sup> The p.Pro555Ala substitution was found in two patients, both carriers of a sarcomeric mutation. A family study in one of the two cases showed that p.Pro555Ala and the MYH7 mutation cosegregated with the disease. The fact that the proband's son was healthy at the age of 28 does not rule out the possibility of developing the disease later on, as his father and aunt were diagnosed at ages 51 and 48, respectively. Interestingly, this variant has been identified as damaging and associated with conotruncal heart defects.<sup>42</sup>

In summary, the rare variants found in the studied GATA genes have the potential to act as modifiers of the clinical phenotype in our cohort of patients, with different associations depending on the presence of sarcomeric mutations. Thus, among patients without known sarcomeric mutations, rare variants in GATA genes were associated with higher thickness of left ventricle walls, although only reached statistical significance in the measure of posterior wall. In addition, patients carrying both rare variants in GATA genes and mutations in sarcomeric genes developed angina more frequently than carriers of sarcomeric mutations alone. This could be explained because angina is a consequence of the insufficiency of blood flow reaching the coronary arteries to the thickened myocardium, the latter is also associated with the presence of GATA variants, and because GATA factors can be critical regulators of cardiac angiogenesis, as it is demonstrated for GATA4.<sup>43</sup>

Several GATA gene mutations have been described in relation to several human diseases which include mainly different types of cancer and congenital heart diseases.<sup>44</sup> All of our patients had left ventricular hypertrophy and none of them had symptoms of these diseases in which GATA variants have been implicated, although we cannot completely rule out that the GATA variants we have found are associated with undiagnosed diseases.

These facts suggest that rare functional GATA variants, although relatively infrequent in HCM, could be modifiers of the phenotypic expression of the disease, adding novel genetic factors to the list of genes associated with the

**Table 4** Effects of GATA rare variants in demographic, clinical, and echocardiographic characteristics of patients with hypertrophic cardiomyopathy

	S	S+G	p Value	NS	NS+G	p Value
BMI (kg/m <sup>2</sup> )	26.8±5	27±4	0.80	27.1±5	26.4±2	0.73
Family history (%)						
Hypertrophic cardiomyopathy	39	66.7	0.22	22.3	27.3	0.79
Sudden cardiac death	19.3	33.3	0.63	2.2	25	0.25
Clinical symptoms (%)						
Angina	23.2	55.5	0.029	34.1	8.3	0.14
Syncope	15.78	11.11	0.72	18.60	8.3	0.50
Dyspnea	62.1	44.4	0.19	55.2	66.7	0.23
Echocardiographic data						
Interventricular septum (mm)	20.9±6	20.7±5	0.98	19±5	21±3	0.11
Posterior wall (mm)	14.9±5	12.9±3	0.40	14±3	16±4	0.019
LVOTO (%)	41.1	55.5	0.37	47.7	33.3	0.61

Quantitative variables are shown as mean±SD.

S, patients carrying sarcomeric mutation; G, patients carrying GATA variant; NS, patients not carrying sarcomeric mutation; p value was obtained by multivariate analysis after adjusting by gender and age.

BMI, body mass index; LVOTO, left ventricular outflow tract obstruction.

phenotypic expression and the severity of hypertrophy, as are polymorphisms of the *MEF2C* gene,<sup>9</sup> renin–angiotensin system,<sup>6</sup> or peroxisome proliferator-activated receptor- $\gamma$  coactivator 1  $\alpha$ ,<sup>4,5</sup> among others. In addition, these results could help in our understanding of the mechanism and provide a novel insight into the etiology and epidemiology of HCM, and furthermore, would contribute to the diagnosis and therapy of this disease.

This study has some limitations. First, despite the ethnic homogeneity of our sample population, the small number of patients requires the replication in larger populations to confirm these findings. Second, the control group was not evaluated by echocardiography to confirm the absence of hypertrophy, and therefore, there could be an underestimation of the impact of our findings. Third, although *in silico* analyses showed putative functional effects, *in vitro* studies are necessary to delineate how these variants affect the expression and activity of these transcription factors.

**Acknowledgements** The authors thank the patients and their families who gave consent to participate in this study.

**Contributors** EC and IR contributed to the conception and design of the work. CA-M, JR-R, MM, JG, MN-D, CM, JBC-A, and IR contributed to the acquisition, analysis, or interpretation of data. CA-M and IR drafted the manuscript. JR-R, MM, JG, EC, MN-D, CM, and JBC-A revised the manuscript critically for important intellectual content. All the authors approved the final version to be published. All the authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding** This work was financed by Plan Nacional de I+D+i 2008–2011; Plan Estatal de I+D+i 2013–2016; Instituto de Salud Carlos III (ISCIII)–Fondo Europeo de Desarrollo Regional (FEDER) (PI07/0659, PI10/00173 and PI13/00497); Plan de Ciencia, Tecnología e Innovación 2013–2017 del Principado de Asturias (GRUPIN14–028); Instituto Reina Sofía de Investigación Nefrológica; Fundación Renal Iñigo Álvarez de Toledo; and Red de Investigación Renal-RedInRen from ISCIII (RD06/0016/1013 RD12/0021/1023 and RD16/0009/0017). CA-M and IR were financially supported by Fundación para el Fomento en Asturias de la Investigación Científica Aplicada y la Tecnología (FICYT).

**Disclaimer** The funders had no involvement in the study design; in the collection, analysis, and interpretation of the data; in the writing of the report; and in the decision to submit the paper for publication.

**Competing interests** None declared.

**Ethics approval** Ethics Committee for Clinical Investigation of Asturias.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Elliott PM, Anastakis A, Borger MA, *et al.* Authors/Task Force members. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). *Eur Heart J* 2014;35:2733–79.
- Thierfelder L, Watkins H, MacRae C, *et al.* Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell* 1994;77:701–12.
- Chiu C, Bagnall RD, Ingles J, *et al.* Mutations in alpha-actinin-2 cause hypertrophic cardiomyopathy: a genome-wide analysis. *J Am Coll Cardiol* 2010;55:1127–35.
- Chiu C, Tebo M, Ingles J, *et al.* Genetic screening of calcium regulation genes in familial hypertrophic cardiomyopathy. *J Mol Cell Cardiol* 2007;43:337–43.
- Chen SN, Czernuszewicz G, Tan Y, *et al.* Human molecular genetic and functional studies identify TRIM63, encoding Muscle RING Finger Protein 1, as a novel gene for human hypertrophic cardiomyopathy. *Circ Res* 2012;111:907–19.
- Coto E, Palacin M, Martin M, *et al.* Functional polymorphisms in genes of the Angiotensin and Serotonin systems and risk of hypertrophic cardiomyopathy: AT1R as a potential modifier. *J Transl Med* 2010;8:64.
- Funada A, Konno T, Fujino N, *et al.* Impact of renin-angiotensin system polymorphisms on development of systolic dysfunction in hypertrophic cardiomyopathy. Evidence from a study of genotyped patients. *Circ J* 2010;74:2674–80.
- Hussain S, Asghar M, Javed Q. Resistin gene promoter region polymorphism and the risk of hypertrophic cardiomyopathy in patients. *Transl Res* 2010;155:142–7.
- Alonso-Montes C, Naves-Diaz M, Fernandez-Martin JL, *et al.* New polymorphisms in human *MEF2C* gene as potential modifier of hypertrophic cardiomyopathy. *Mol Biol Rep* 2012;39:8777–85.
- Marian AJ. Modifier genes for hypertrophic cardiomyopathy. *Curr Opin Cardiol* 2002;17:242–52.
- Xu J, Li Z, Ren X, *et al.* Investigation of pathogenic genes in Chinese sporadic hypertrophic cardiomyopathy patients by whole exome sequencing. *Sci Rep* 2015;5:16609.
- Akazawa H, Komuro I. Roles of cardiac transcription factors in cardiac hypertrophy. *Circ Res* 2003;92:1079–88.
- Orkin SH. GATA-binding transcription factors in hematopoietic cells. *Blood* 1992;80:575–81.
- Dorfman DM, Wilson DB, Bruns GA, *et al.* Human transcription factor GATA-2. Evidence for regulation of preproendothelin-1 gene expression in endothelial cells. *J Biol Chem* 1992;267:1279–85.



- 15 Kawana M, Lee ME, Quertermous EE, *et al.* Cooperative interaction of GATA-2 and AP1 regulates transcription of the endothelin-1 gene. *Mol Cell Biol* 1995;15:4225–31.
- 16 Bupha-Intr T, Haizlip KM, Janssen PM. Role of endothelin in the induction of cardiac hypertrophy in vitro. *PLoS ONE* 2012;7:e43179.
- 17 Molkentin JD. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. *J Biol Chem* 2000;275:38949–52.
- 18 Liang Q, De Windt LJ, Witt SA, *et al.* The transcription factors GATA4 and GATA6 regulate cardiomyocyte hypertrophy in vitro and in vivo. *J Biol Chem* 2001;276:30245–53.
- 19 van Berlo JH, Elrod JW, van den Hoogenhof MM, *et al.* The transcription factor GATA-6 regulates pathological cardiac hypertrophy. *Circ Res* 2010;107:1032–40.
- 20 Maron BJ, McKenna WJ, Danielson GK, *et al.*, American College of Cardiology Foundation Task Force on Clinical Expert Consensus Documents; European Society of Cardiology Committee for Practice Guidelines. American College of Cardiology/European Society of Cardiology Clinical Expert Consensus Document on Hypertrophic Cardiomyopathy. A report of the American College of Cardiology Foundation Task Force on Clinical Expert Consensus Documents and the European Society of Cardiology Committee for Practice Guidelines. *Eur Heart J* 2003;24:1965–91.
- 21 Garcia-Castro M, Coto E, Reguero JR, *et al.* [Mutations in sarcomeric genes MYH7, MYBPC3, TNNT2, TNNI3, and TPM1 in patients with hypertrophic cardiomyopathy]. *Rev Esp Cardiol* 2009;62:48–56.
- 22 Gomez J, Reguero JR, Moris C, *et al.* Mutation analysis of the main hypertrophic cardiomyopathy genes using multiplex amplification and semiconductor next-generation sequencing. *Circ J* 2014;78:2963–71.
- 23 Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16:1215.
- 24 Rodríguez I, Coto E, Reguero JR, *et al.* Role of the CDKN1A/p21, CDKN1C/p57, and CDKN2A/p16 genes in the risk of atherosclerosis and myocardial infarction. *Cell Cycle* 2007;6:620–5.
- 25 Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomarkers* 2010;14:533–7.
- 26 Barrett JC, Fry B, Maller J, *et al.* Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5.
- 27 Gabriel SB, Schaffner SF, Nguyen H, *et al.* The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–9.
- 28 Komati H, Maharsy W, Beauregard J, *et al.* ZFP260 is an inducer of cardiac hypertrophy and a nuclear mediator of endothelin-1 signaling. *J Biol Chem* 2011;286:1508–16.
- 29 Zhang X, Azhar G, Chai J, *et al.* Cardiomyopathy in transgenic mice with cardiac-specific overexpression of serum response factor. *Am J Physiol Heart Circ Physiol* 2001;280:H1782–92.
- 30 McKinsey TA, Olson EN. Toward transcriptional therapies for the failing heart: chemical screens to modulate genes. *J Clin Invest* 2005;115:538–46.
- 31 Li RG, Li L, Qiu XB, *et al.* GATA4 loss-of-function mutation underlies familial dilated cardiomyopathy. *Biochem Biophys Res Commun* 2013;439:591–6.
- 32 Xu L, Zhao L, Yuan F, *et al.* GATA6 loss-of-function mutations contribute to familial dilated cardiomyopathy. *Int J Mol Med* 2014;34:1315–22.
- 33 Abecasis GR, Auton A, Brooks LD, *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56–65.
- 34 Fisch S, Gray S, Heymans S, *et al.* Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy. *Proc Natl Acad Sci USA* 2007;104:7074–9.
- 35 Buratti E, Baralle FE. Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* 2004;24:10505–14.
- 36 van Berlo JH, Elrod JW, Aronow BJ, *et al.* Serine 105 phosphorylation of transcription factor GATA4 is necessary for stress-induced cardiac hypertrophy in vivo. *Proc Natl Acad Sci USA* 2011;108:12331–6.
- 37 Charron F, Paradis P, Bronchain O, *et al.* Cooperative interaction between GATA-4 and GATA-6 regulates myocardial gene expression. *Mol Cell Biol* 1999;19:4355–65.
- 38 Morin S, Charron F, Robitaille L, *et al.* GATA-dependent recruitment of MEF2 proteins to target promoters. *Embo J* 2000;19:2046–55.
- 39 Wang J, Sun YM, Yang YQ. Mutation spectrum of the GATA4 gene in patients with idiopathic atrial fibrillation. *Mol Biol Rep* 2012;39:8127–35.
- 40 Poirier O, Nicaud V, McDonagh T, *et al.* Polymorphisms of genes of the cardiac calcineurin pathway and cardiac hypertrophy. *Eur J Hum Genet* 2003;11:659–64.
- 41 Laval F, Bessonnard S, Ohnishi Y, *et al.* Bmi1 facilitates primitive endoderm formation by stabilizing Gata6 during early mouse development. *Genes Dev* 2012;26:1445–58.
- 42 Osoegawa K, Parodi C, Mohammed N, *et al.* Sequencing candidate genes associated with conotruncal heart defects. *Presented at the 63rd Annual Meeting of the American Society of Human Genetics*, Boston, MA (Abstract). 2013.
- 43 Heineke J, Auger-Messier M, Xu J, *et al.* Cardiomyocyte GATA4 functions as a stress-responsive regulator of angiogenesis in the murine heart. *J Clin Invest* 2007;117:3198–210.
- 44 Lentjes MH, Niessen HE, Akiyama Y, *et al.* The emerging role of GATA transcription factors in development and disease. *Expert Rev Mol Med* 2016;18:e3.
- 45 Wang S, Fu C, Wang H, *et al.* Polymorphisms of the peroxisome proliferator-activated receptor-gamma coactivator-1alpha gene are associated with hypertrophic cardiomyopathy and not with hypertension hypertrophy. *Clin Chem Lab Med* 2007;45:962–7.