Genetics of coarctation of the aorta in Iceland

Þorsteinn Björnsson

Thesis for the degree of Bachelor of Science
University of Iceland
Faculty of Medicine
School of Health
Genetics of coarctation of the aorta in Iceland

Porsteinn Björnsson

Thesis for the degree of Bachelor of Science
Supervisor: Hilma Hólm
Co-supervisors: Tómas Guðbjartsson, Hróðmar Helgason, Daniel F. Guðbjartsson, Unnur Porsteinsdóttir, Kári Stefánsson

Faculty of Medicine, University of Iceland, deCODE genetics, Landspitali, The National University Hospital, School of Engineering and Natural Sciences, University of Iceland
January 2016
Abstract

Introduction: Coarctation of the aorta (CoA) accounts for 3.8% of all congenital heart disease in Iceland. Despite excellent surgical outcomes, CoA can be a life-long disease with high rates of long-term cardiovascular complications. The underlying genetic basis and pathogenesis of CoA remains largely unknown. The aim of this study was to search for sequence variants that affect the risk of CoA in Iceland.

Methods: The CoA cases were Icelanders (N=132) who received the discharge diagnosis of CoA at Landspitali, The National University Hospital (LUH) in Reykjavik between 1984 and 2015. Detailed phenotypic information on CoA cases was gathered through a centralized electronic database on patient’s records (Saga system) as well as paper records at LUH. To identify sequence variants that associate with CoA risk, genome-wide association analysis (GWAS) was performed using 25.5 million sequence variants identified through whole-genome sequencing of 8,453 Icelanders that were subsequently imputed into a large fraction of Icelanders. The GWAS was performed, with the 132 CoA cases and as controls 339,213 Icelanders without CoA, using logistic regression, adjusting for gender, age and county of origin.

Results: Through the CoA GWAS analysis we identified a rare (0.35%) missense variant c.2161C>T in exon 18 of the MYH6 gene that associates with increased risk of CoA. MYH6 is a large gene that encodes the alpha myosin heavy chain (αMHC), a major component of the sarcomere of cardiac muscle. The c.2161C>T mutation associates with CoA with large effect, an odds ratio of 31.4 (95% confidence interval; 14.08, 69.83) and with high significance, P of 3.3 x 10^{-17}. The c.2161C>T results in a change of arginine to tryptophan at amino acid 721 (p.Arg721Trp) in the converter domain of the αMHC protein. This same mutation has previously been reported to associate with sick sinus syndrome. Of the 132 CoA cases, 24 were carriers of c.2161C>T; no significant phenotypic difference was found between CoA carriers and non-carriers of c.2161C>T. This may in part be explained by the small size of the study. The MYH6 c.2161C>T was not found outside of Iceland.

Conclusions and Discussion: The MYH6 gene has not previously been reported to associate with CoA although other very rare mutations in MYH6 have been linked to both familial hypertrophic cardiomyopathy and familial atrial septal defect. The c.2161C>T mutation explains a large fraction or 19% of CoA cases in Iceland, a figure rarely reported in genetic studies of congenital heart disease. Expression of MYH6 has not been detected in the aorta but it is highly expressed throughout life in the atrium and in the ventricle during embryonic cardiogenesis. The p.Arg721Trp mutation in the converter domain of αMHC is predicted to be damaging and might thus affect the contractile function of αMHC in the heart. It is conceivable that p.Arg721Trp predisposes to CoA by reducing the contraction of the developing heart thus reducing blood flow through the aorta which is in line with the hemodynamic theory, a leading theory of CoA pathogenesis. This hypothesis is compatible with the fact that MYH6 is not expressed in the aorta.
Acknowledgements

I would like to thank my supervisors Hilma Hólm, MD, Tómas Guðbjartsson, MD, PhD, Hróðmar Helgason, MD, Daniel F. Guðbjartsson, PhD, Unnur Þorsteinsdóttir, PhD and Kári Stefánsson, MD, PhD for exceptional guidance, support, encouragement and patience. I would also like to thank Ingibjörg Richter, Systems Analyst at Landspítali University Hospital, Gunnhildur Jóhannsdóttir, Office Manager at Landspítali University Hospital and Sverrir Gunnarsson, Cardiovascular Medicine Fellow at University of Wisconsin School of Medicine and Public Health, for their valuable assistance. This journey has been a valuable experience and I thank all above for making this possible.
Table of contents

Abstract ................................................................................................................................. 4
Acknowledgements ............................................................................................................. 5
Table of contents ............................................................................................................... 6
List of figures ..................................................................................................................... 7
List of tables ...................................................................................................................... 7
Abbreviations .................................................................................................................... 8
1. Introduction .................................................................................................................. 9
  1.1 Congenital heart disease and coarctation of the aorta ........................................ 9
  1.2 Pathogenesis of coarctation of the aorta ............................................................ 10
  1.3 Genetics of congenital heart disease ................................................................. 11
  1.4 Genetics of coarctation of the aorta ................................................................. 12
  1.5 Whole–genome sequencing to search for rare variants that associate with disease ... 13
  1.6 Aim of study ...................................................................................................... 15
2. Methods ....................................................................................................................... 16
  2.1 Identification and phenotyping of coarctation of the aorta cases .................... 16
  2.2 Genealogy of Icelanders, the Book of Icelanders .............................................. 16
  2.3 Genotyping, imputation and whole-genome sequencing ..................................... 17
  2.4 Association analysis ......................................................................................... 18
  2.5 Phenotypic differences between carriers and non-carriers of c.2161C>T ............ 18
3. Results ......................................................................................................................... 19
  3.1 Coarctation of the aorta sample set ................................................................... 19
  3.2 Genome-wide association analysis .................................................................... 20
  3.3 Phenotypic differences between carriers and non-carriers of c.2161C>T ............ 23
  3.4 Association of c.2161C>T with other cardiac conditions .................................... 24
4. Conclusions and Discussion ....................................................................................... 26
  4.1 Coarctation of the aorta in Iceland .................................................................... 26
  4.2 MYH6 c.2161C>T explains a large fraction of coarctation of the aorta in Iceland .... 27
  4.3 Pathogenesis of c.2161C>T in coarctation of the aorta ....................................... 27
  4.4 c.2161C>T in other cardiac conditions ............................................................ 28
  4.5 Future studies ..................................................................................................... 28
5. References .................................................................................................................... 29
List of figures

Figure 1. Coarctation of the aorta ................................................................. 9
Figure 2. Initial diagnoses of coarctation of the aorta cases .................................................. 19
Figure 3. Manhattan plot of coarctation of the aorta genome-wide association study ............. 21
Figure 4. Region plot for the association of variants with coarctation of the aorta ............... 21
Figure 5. Transcript and protein structures of MYH6 ....................................................... 22
Figure 6. A 3D structural model of chicken smooth muscle myosin that is homologous to MHCα.... 22

List of tables

Table 1. Frequencies of various phenotypes among coarctation of the aorta cases .................... 20
Table 2. Frequencies of various phenotypes among non-carriers and carriers of c.2161C>T ..... 24
Table 3. Association of c.2161C>T with various cardiac conditions ........................................ 25
Abbreviations

CoA  Coarctation of the aorta
CHD  Congenital heart disease
DA   Ductus arteriosus
TH   Tubular hypoplasia
LVOTO Left ventricular outflow tract obstruction
BAV  Bicuspid aortic valve
VSD  Ventricular septal defect
TS   Turner syndrome
NGS  Next-generation sequencing
WGS  Whole-genome sequencing
SNP  Single-nucleotide polymorphism
INDEL Insertion/deletion
GWAS Genome-wide association analysis
LUH  Landspitali, The National University Hospital
OR   Odds ratio
CI   Confidence interval
αMHC Alpha myosin heavy chain
ATP  Adenosine triphosphate
ASD  Atrial septal defect
SSS  Sick sinus syndrome
PM   Pacemaker implantation
AF   Atrial fibrillation
DCM  Dilated cardiomyopathy
HCM  Hypertrophic cardiomyopathy
AS   Aortic stenosis
βMHC Beta myosin heavy chain
IPSC Induced pluripotent stem cell
1. Introduction

1.1 Congenital heart disease and coarctation of the aorta

The impact of congenital heart disease (CHD) on global health is profound. Worldwide, 1.35 million infants are born with CHD each year. CHD is the most common type of birth defect and the leading cause of birth defect-related infant morbidity and mortality (1). CHD is also identified in 10% of stillbirths and thought to play a major role in early fetal death (2). Advances in the management of patients with CHD have enabled most patients to reach adulthood. Adults now constitute the majority of the CHD population. Adult CHD is associated with high rates of cardiac comorbidities and may require life-long cardiology follow up (3).

Coarctation of the aorta (CoA) is the most common birth defect of the aorta (4). In Iceland, it accounts for 3.8% of all CHD with an incidence of about 1 per 1,500 live births (5). CoA is a narrowing of the aorta. The site of CoA is usually at or near the point of the insertion of the ductus arteriosus (DA), just distal to the origin of the left subclavian artery (LSCA) (Figure 1). From a histological perspective, the lesion is a tissue shelf in the posterolateral aortic wall and consists mainly of infolding and thickening of the aortic media (6). A neonatal presentation is more often associated with a shelf plus aortic arch and isthmic hypoplasia (tubular hypoplasia; TH), whereas with a later presentation, these areas are larger. In a classification system of CHD, based on similarity, complexity and suspected embryologic origin of defects, CoA is classified as a left ventricular outflow tract obstruction (LVOTO). LVOTO type CHD include bicuspid aortic valve (BAV), congenital aortic stenosis, interrupted aortic arch type A, Shone syndrome and hypoplastic left heart syndrome (7).

Figure 1. Coarctation of the aorta.
The site of CoA is usually at or near the point of the insertion of the ductus arteriosus (DA), just distal to the origin of the left subclavian artery (LSCA). Adapted from The PedHeart Community Web Courtesy of Scientific Software Solutions, Inc. (8). Adapted with permission.

CoA may occur as an isolated defect or with other CHD. Almost any CHD may associate with CoA. In general, there is a strong association between CoA and other LVOTOs. Indeed, the most common association is with BAV, occurring in about 50% of patients with CoA. CoA is also an important component of hypoplastic left heart and Shone syndrome, rare congenital left
heart disorders with significant morbidity and mortality. The second most common association is with ventricular septal defect (VSD), occurring in about 20% of patients (9, 10). Intracranial aneurysms are identified in up to 10% of adults with CoA (11). Extracardiac birth defects occur in about 25% of patients with CoA, some of which have a recognized syndrome (12).

The clinical presentation of CoA depends on the severity of obstruction to left-ventricular outflow. The severe form of CoA presents in neonates as congestive heart failure or cardiogenic shock and requires immediate and definitive medical and surgical treatment. On the other hand, CoA may present later in childhood and even well into adulthood. These patients are usually asymptomatic but may present with symptoms related to hypertension (e.g. frequent headaches) or reduced blood flow to lower limbs (e.g. lower-limb fatigue with exercise) (13). A rare but severe initial presenting feature in adults is rupture of an intracranial aneurysm (11). Physical findings in CoA include heart murmurs or discrepant arterial pulses and systolic blood pressures in the upper and lower limbs. The diagnosis of CoA is usually confirmed by echocardiography. Prenatal diagnosis is possible but challenging (14).

The natural history of untreated CoA is poor and intervention is indicated in essentially all patients with CoA (15). In Iceland and its neighboring countries, surgery is the treatment of choice but balloon angioplasty with or without stenting has become an alternative option in selected patients. The most common surgical procedure is resection of the coarctation with an end-to-end anastomosis of the aorta. With the exception of patients with complex CHD, outcomes are very good in regard to short-term survival, elimination of the obstruction and normalization of blood pressure (16-18). Overall, following successful repair of CoA, survival into adulthood is expected but patients remain at high risk for late cardiovascular complications and have a shorter life expectancy (19, 20). Hypertension is a major complication, occurring in about 10 to 40% of patients at rest and at higher rates during exercise (21-23). Other complications include recoarctation (24), complications of associated CHD (25), cerebrovascular accidents (26) and aortic aneurysm/dissection/rupture (27). The most common causes of death have been reported to be premature coronary artery disease, heart failure and sudden death (20). As such, in many cases, CoA is a life-long disease requiring continuous cardiology follow-up.

1.2 Pathogenesis of coarctation of the aorta

The heart is the first functional organ to form in developing embryos. Disruptions of cardiac development cause CHD. Cardiac development is a complex morphogenetic event involving multiple cell lineages controlled by overlapping genetic regulatory networks and hemodynamic influences, which we are only beginning to understand. Because of this, the precise pathogenesis of CoA is poorly understood.

The aortic arch and its branches develop during the 6th to 8th week of human embryogenesis. The thoracic aortic arch and isthmus derive from the left 4th aortic arch and the DA derives from a distal portion of the left 6th aortic arch. Thus, disruptions to left 4th and 6th aortic arch development cause CoA (28). Two theories have gained the widest recognition in explaining the mechanism by which CoA is produced, the hemodynamic theory and the ductal tissue theory.
The hemodynamic theory is based on the assumption that vessel diameter is proportional to fetal blood flow. In the normal fetus the blood flow across the aortic isthmus is lower than the blood flow across the ascending and descending aorta. This is reflected in a smaller diameter of the isthmus than of the ascending and descending aorta. A similar relationship between blood flow and aortic diameter has been shown in various CHD with increased or decreased aortic outflow (29). According to the hemodynamic theory, lesions that result in decreased left ventricular outflow promote development of CoA by reducing blood flow through the aortic isthmus (30). The theory is consistent with the following: (1) CoA is commonly associated with TH, LVOTOs and VSDs; (2) TH is the most definitive antenatal sign of postnatal CoA (31); and (3) almost any CHD may associate with CoA but pulmonary stenosis, pulmonary atresia and tetralogy of Fallot, lesions that increase aortic isthmus flow, are extremely rare (9). The theory fails to explain the occurrence of CoA without lesions that reduce left ventricular outflow, and vice versa. Interestingly, the theory may explain the fact that CoA is more common in Turner syndrome when a web neck is involved (32). A web neck occurs secondary to lymphatic obstruction, which may lead to distended thoracic ducts compressing the fetal ascending aorta and thereby promote development of CoA.

The ductal tissue theory suggests that postnatal constriction of aberrant ductal tissue extending into the aortic wall results in CoA (33). It is consistent with the following: (1) ductal tissue has been shown to make up the inner part of the coarctation lesion (34); (2) rather than being ectopic, the ductal tissue in the aortic wall is thought to represent the original distal wall of the left sixth aortic arch (35); (3) CoA often becomes manifest after ductal closure and patients may benefit from prostaglandin E1 promoting ductal patency. The theory fails to explain the occurrence of CoA in other locations such as in the abdominal aorta and the common association of TH with CoA.

CoA is strongly associated with other vascular diseases including BAV, intracranial aneurysms and late vascular complications such as hypertension. This has prompted some to argue that CoA should be regarded as part of diffuse arteriopathy, rather than just being a local vascular stenosis (36-38). However, it remains difficult to determine the balance between genetic and adaptive factors causing these vascular diseases. An interesting approach supporting the assumption of CoA being part of a diffuse arteriopathy is that CoA is caused by a developmental abnormality of neural crest tissue. It would be consistent with the fact that the neural crest plays an important role in development of the left ventricular outflow tract, aortic arch and cervicocephalic arteries. The common association of head-and-neck abnormalities with CoA, particularly when CoA is with BAV, further strengthens this approach (39).

1.3 Genetics of congenital heart disease

Epidemiologic studies suggest that CHD arises primarily through genetic abnormalities but environmental factors should not be underestimated. Targeted gene deletion studies in experimental models have revealed over 500 genes that can induce CHD when mutated (40). It is conceivable that a similar number of human CHD disease genes exist. However, the genetic basis of CHD in humans is poorly understood (41). This is because the genetic methods used to identify causal genes (e.g. linkage analysis) have, until recently, relied on familial forms of CHD with high penetrance. Owing to the lack of large families with multiple affected individuals (rare in CHD) these attempts have been, with some
exceptions (42-44), largely unsuccessful. With the availability of only small families and sporadic cases, investigators have relied on screening candidate genes in CHD samples. The great majority of genes identified with this approach have not been validated by functional studies nor implicated in multiple independent cases, and, thus, their causality in CHD remains unclear.

Along with long-recognized chromosomal abnormalities, these approaches have yielded over 50 human disease genes implicated in CHD that explain minority of cases of CHD (40). Importantly, these studies have also provided the following insights (45): (1) CHD is genetically heterogeneous; (2) cardiac developmental genes are the major players in CHD; (3) each CHD mutation may give rise to a variety of CHD phenotypes, suggesting interactions of the identified mutation with genomic context and/or the environment; (4) CHD mutations commonly alter protein dosage.

With the advent of high-throughput next-generation sequencing (NGS) technologies that allows the whole-genomes of individuals to be sequenced, discoveries of the genetic causes of CHD are accelerating. Importantly, since whole-genome sequencing (WGS) can detect a large fraction of the genomic variation of an individual, it allows the use of non-Mendelian or sporadic cases (the most common presentation of CHD) to search for rare variants in causal genes (46).

It is important to uncover the genetic basis of CHD in order to understand the pathogenesis of CHD and estimate disease risk, both vital elements for disease prevention. From a clinical standpoint, it is important for the caring physician to determine whether there is an underlying genetic cause for various reasons. First, from a psychosocial perspective for the patient and family. Second, there may be extracardiac organ involvement and prognostic information for clinical outcomes. Third, there may be reproductive risks for the patient and family and genetic testing may be appropriate for other family members.

1.4 Genetics of coarctation of the aorta

Genetic studies of CoA are usually carried out along with other LVOTOs (see 1.1 above). A strong genetic component has been established that is likely shared by these malformations as they tend to cosegregate in families (47). The evidence for a strong genetic component of LVOTOs comes from their long-recognized association with genetic syndromes (48). Furthermore, inheritance studies of LVOTOs have demonstrated a high relative risk for first degree relatives (36.9) and high heritability estimates (between 0.60-0.90) (49).

The syndrome with the strongest association with CoA is Turner syndrome (TS) a condition in females where one X chromosome or a part of it is absent. The most common CHD in TS are CoA and BAV but also other LVOTOs. TS accounts for about 5 percent of CoA in females (50). Interestingly, CoA and other LVOTOs are about two times more common in males than in females (51). Moreover, Turner patients suffer from excess cardiovascular morbidity and premature mortality associated with similar complications as CoA (e.g. hypertension, premature coronary artery disease). This has prompted some to investigate the precise genetic link between the risk of developing LVOTOs and having one X chromosome. However, the link is not well understood (37). William-Beuren syndrome is caused by a 1.5 to 1.8 Mb hemizygous deletion on chromosome 7q11.23, an area that encompasses 28 genes.
Supravalvar AS is by far the most common CHD associated with this syndrome but other LVOTOs such as CoA may occur. The cardiovascular phenotype of William-Beuren syndrome is caused by haploinsufficiency of the ELN gene, identified in at least 70 percent of cases (48). LVOTOs can be associated with Jacobsen syndrome (JS) which is caused by 11q terminal deletions between 5-20 Mb. A transcription factor, ETS-1, within this region has been implicated in LVOTOs (52). LVOTOs can be associated with Noonan syndrome which is caused by mutations in genes encoding proteins that are part of the Ras/Raf/MEK/ERK signaling pathway. Mutations in PTPN11 are the most common and explain about 50 percent of the Noonan syndrome cases. In monogenic syndromes, mutations in genes such as TBX5 in Holt-Oram syndrome, DHCR7 in Smith-Lemli-Opitz syndrome, and ZIC3 in X-linked heterotaxy, have been linked with LVOTOs (48).

The vast majority of cases of LVOTOs are non-syndromic. Mutations in NOTCH1, initially found through linkage analysis of families with aortic valve disease, are found across various LVOTOs (42). NOTCH1 encodes a receptor in a developmentally important signaling pathway. Two specific mutations that reduce ligand (JAGGED1) induced NOTCH1 signaling have been identified, suggesting that the levels of NOTCH1 signaling are tightly regulated during cardiovascular development and that relatively minor alterations in the signaling may induce LVOTOs (53). In addition to copy number variations (CNVs) in the above described syndromes, CNVs have also been described in non-syndromic LVOTOs. These include deletions in 16q24 that affect a FOX gene cluster (FOXF1, FOXL1, and FOXC1), a deletion at 6q24 harboring the MAPK signaling cofactor TAB2 shown to be responsible for the LVOT defect, and terminal deletions of 15q26 that harbor MCTP2 gene that is a causal gene for LVOTOs. The CNVs associated with syndromic and/or sporadic LVOTOs usually harbor many genes. The causal LVOTO genes within these CNVs have mainly been identified using model systems like the zebra fish or the mouse (54). The translocation 46,XY,t(1;5)(p36.11;q31.2) that is associated with pervasive developmental delay and LVOTOs disrupts two genes, AHDC1 and MATR3. Through gene inactivation studies in mouse, MATR3 was found to induce LVOTOs but not AHDC1 (55).

Although many genes have been linked to the development of LVOTOs, in the vast majority of cases the genetic cause is not identified. Thus, a substantial fraction of the heritability of CoA and other LVOTOs is unexplained.

1.5 Whole-genome sequencing to search for rare variants that associate with disease

As discussed above, NGS technologies have provided the means to sequence the whole genomes of a large number of individuals, allowing for the detection of low frequency (freq. <5% to >1%) and rare variants (freq. <1%) that can be tested for association to diseases and other traits. These variants are single-nucleotide polymorphisms (SNPs) or insertions/deletions (INDELs) (56-58). Previously, deCODE scientists have performed genome-wide association analysis (GWAS) based on variants identified through WGS of Icelanders that was followed by estimation of their genotype probabilities in the Icelandic population using imputation assisted by long range phasing of haplotypes (59, 60). This uncovered associations of many low frequency and rare sequence variants with a large number of traits (61-67). Icelanders have experienced more genetic drift than most other populations used in GWAS (68, 69). In particular, this has affected the frequency spectrum of variants that are rare in neighboring
outbred populations. While some rare variants have been lost from the Icelandic population, the frequency of others have increased by founder events, leading to greater power for detection in phenotype association tests. Furthermore, because a large fraction of Icelanders (~150,000) have been genotyped on SNP arrays at deCODE, the genomes of a large fraction of Icelanders has been long-range phased. This phasing has made high accuracy imputations possible for variants with frequency as low as 0.03%. This allows GWAS to be performed at deCODE with sequence variants identified through WGS for diseases where cases and controls have not been directly whole-genome sequenced themselves, but are imputed with sequence variants found through WGS of other Icelandic samples. Here we use this dataset to identify sequence variants in the human genome that affect the risk of CoA.
1.6 Aim of study

The aim of this study is divided into two parts:

a) To use the genetic database at deCODE to identify rare sequence variants that affect the risk of developing CoA in Iceland

b) To determine if there is any phenotypic difference between non-carriers and carriers of the identified sequence variants
2. Methods

2.1 Identification and phenotyping of CoA cases

The CoA sample set includes Icelanders (population of about 330,000) who received the discharge diagnosis of CoA at Landspitali, The National University Hospital (LUH) in Reykjavik between 1984 and 2015 (06.03.2015). LUH is a tertiary referral hospital and the only hospital in Iceland with pediatric cardiologists. All CoA cases in Iceland are referred to LUH. CoA cases were identified either through diagnosis codes of CoA (ICD-9 code 747.1, ICD-10 code Q25.1) registered between 1990 and 2015 or procedure codes of CoA (WHO codes 1-273, 5-369, 5-382 and 5-387, NOMESCO codes FDJ 00, FDJ 10, FDJ 20, FDJ 30, FDJ 42 and FDJ 96) registered between 1984 and 2015. CoA was defined as a congenital narrowing of aorta, the diagnosis of which was confirmed by echocardiography and/or cardiac catheterization by a cardiologist. Stillbirths, early fetal deaths and non-Icelanders were excluded from the study.

Through a centralized electronic database on patient’s records (Saga system) as well as paper records at LUH, the diagnosis of CoA was confirmed and detailed phenotypic information on each CoA case was gathered. Special emphasis was put on associated CHD, late cardiovascular complications and the classification of CoA cases into six different types. Mild CoA was defined as untreated CoA. Moderate CoA was defined as CoA receiving treatment but not presenting with severe disease. Severe CoA was defined as CoA presenting as congestive heart failure or cardiogenic shock without the presence of major complicating CHD. CoA with VSD was defined as CoA with single or multiple large VSD/s presenting as congestive heart failure or cardiogenic shock. CoA with complex CHD was defined as CoA with complex CHD presenting as heart failure or cardiogenic shock. Atypical CoA was defined as coarctation of the abdominal aorta or interrupted aortic arch type A. Information on date of birth and gender came from the Book of Icelanders (contains genealogical information about Icelanders). The phenotypic information was used for phenotype-genotype correlation of CoA cases. Data was filed in Microsoft Office Excel 2013.

The study was approved by the National Bioethics Committee of Iceland. Study approval numbers were VSN-15-053, VSN-15-016, VSN-15-056, VSN-15-058, VSN-15-114, VSN-15-057 and 10-009-S1. Written informed consent was obtained from all study participants. Personal identities were encrypted by a third party system provided by the Icelandic Data Protection Authority before analysis at deCODE.

2.2 Genealogy of Icelanders, the Book of Icelanders

The Book of Icelanders (61) contains 819,410 individuals dating back to 740 AD. Of the 471,284 Icelanders recorded to have been born in the 20th century, 91.1% had a recorded father and 93.7% had a recorded mother in the database. Similarly, of the 183,896 Icelanders recorded to have been born in the 19th century, 97.5% had a recorded father and 97.8% had a recorded mother. The Icelandic genealogy was extracted from many sources. Primarily from church books, censuses, Registers Iceland, local records of inhabitants and other official documents, but also from other sources such as old manuscripts, letters, annals, books of Althingi, books of judgments, books of family pedigrees, registers of farmers, registers of professionals and lists of descendants. The church books and several censuses
have been computerized. The genealogical database was primarily based on the censuses of 1703, 1801 and 1910, but other censuses that have been computerized are from: 1729, 1785, 1816, 1835, 1845, 1860, 1870, 1880, 1890, 1901 and 1930.

2.3 Genotyping, imputation and whole-genome sequencing,

For chip genotyping, 150,656 samples were typed with the Illumina HumanHap300, HumanCNV370, HumanHap610, HumanHap1M, HumanHap660, Omni-1, Omni 2.5 or Omni Express bead chips at deCODE genetics. Chip SNPs were excluded if they had (1) yield less than 95%, (2) minor allele frequency less than 1% in the population or (3) significant deviation from Hardy-Weinberg equilibrium ($P$<0.001), (4) if they produced an excessive inheritance error rate (over 0.001) and (5) if there was substantial difference in allele frequency between chip types (from just a single chip if that resolved all differences, but from all chips otherwise). All samples with a call rate below 97% were excluded from the analysis. The final SNP set used for long-range phasing comprised 676,913 autosomal SNPs.

Long range phasing of all chip-genotyped individuals was performed with methods described previously (59, 60). In brief, phasing is achieved by using an iterative algorithm which phases a single proband at a time given the available phasing information about everyone else that shares a long haplotype identically by state with the proband. Given the large fraction of the Icelandic population that has been chip-typed, accurate long range phasing is available genome-wide for all chip-typed Icelanders. For long range phased haplotype association analysis, we then partitioned the genome into non-overlapping fixed 0.3 cM bins. Within each bin, we observed the haplotype diversity described by the combination of all chip-typed markers in the bin.

The whole genomes of 8,453 Icelanders were sequenced using Illumina technology to a mean depth of at least 10X (median 32X). This dataset contains samples obtained using three different library preparation methods from Illumina. In addition sequencing was performed using three different types of Illumina sequencing instruments.

a) Standard TruSeq DNA library preparation method. Illumina sequencers GAIIx and/or HiSeq 2000 sequencers.

b) TruSeq DNA PCR-free library preparation method. Illumina HiSeq 2500 sequencers.

c) TruSeq Nano DNA library preparation method. Illumina HiSeq X

In the sequencing dataset SNPs and INDELS were identified and genotypes called using joint calling with the Genome Analysis Toolkit HaplotypeCaller (GATK version 3.3.0) (70). Genotype calls were improved by using information about haplotype sharing, taking advantage of the fact that all the sequenced individuals had also been chip-typed and long-range phased.

The sequence variants identified in the 8,453 sequenced Icelanders were then imputed into 150,656 Icelanders who had been genotyped with various Illumina SNP chips and their genotypes phased using long-range phasing (59, 60). The imputation into the chip typed long-range phased individuals was performed with the same model as used by IMPUTE (71). Using genealogic information, from the Book of Icelanders, the sequence variants were imputed into 294,212 first and second-degree relatives of
array genotyped individuals to further increase the sample size for association analysis and the power to detect associations. We identified a total of 25.5 million high quality sequence variants (20,601,114 SNPs and 4,876,374 short INDELs; all with imputation information >0.8; all variants mapped in build hg38) and tested for association with CoA under the multiplicative model.

2.4 Association analysis

The number of chip-typed CoA cases was 38 and the remaining 94 were first or second degree relatives of the 150,656 chip-typed individuals and were not chip-typed themselves but were imputed using genealogical imputation (59, 60). The number of affected males was 82 and affected females was 50. Association testing for case-control analysis was performed using logistic regression, adjusting for gender, age and county of origin. A total of 25.5 million variants were used in the association analysis under a multiplicative model.

To account for inflation in test statistics due to cryptic relatedness and stratification, we applied the method of LD score regression (72). With a set of 1.1M variants we regressed the \( \chi^2 \) statistics from our GWAS scan against LD score and used the intercept as a correction factor. The LD scores were downloaded from a LD score database (61). The estimated correction factor was 1.04 for the multiplicative model of the CoA association.

To correct for multiple testing we used the weighted Holm-Bonferroni method (73) to allocate family wise error rate of 0.05 equally between four annotation-based classes of sequence variants (74). For the multiplicative model, this yielded significance thresholds of \( 3.3 \times 10^{-7} \) for high-impact variants (including stop-gained, frameshift, splice acceptor or donor, \( N = 5,969 \)), \( 6.5 \times 10^{-8} \) for moderate-impact variants (including missense, splice-region variants and in-frame INDELs, \( N = 118,721 \)) \( 5.9 \times 10^{-9} \) for low impact variants (including synonymous variants 3’ and 5’ UTR variants, \( N = 1,797,313 \)) \( 3.0 \times 10^{-9} \) for intergenic and deep intronic variants overlapping DNase hypersensitive sites (\( N = 3,302,617 \)) and for other variants \( 9.9 \times 10^{-10} \) (intergenic and deep intronic, \( N = 20,252,868 \)) (75).

Association of c.2161C>T with other cardiac or cardiac related diseases was also tested, including hypertrophic cardiomyopathy, congenital heart disease, sick sinus syndrome, pacemaker implantation, atrial fibrillation, heart failure, sudden cardiac death, coronary artery disease and ischemic stroke. The controls used in the various case-control analyses of this study consisted of disease-free controls randomly drawn from other genetic studies at deCODE.

2.5 Phenotypic differences between carriers and non-carriers of c.2161C>T

To analyze if there is any phenotypic difference between non-carriers and carriers of the c.2161C>T mutation, we evaluated the frequencies of various phenotypes among the two groups of CoA cases (Table 1). Student’s \( t \)-test was used to test for significant difference in the mean frequency of the variants between non-carriers and carriers, and the odds ratio (OR) was calculated as \( \frac{pa}{1-pa} / \frac{pc}{1-pc} \), where \( pa \) and \( pc \) are the mean frequencies of the variants in non-carriers and carriers, respectively.
3. Results

3.1 Coarctation of the aorta sample set

We identified 136 individuals who received a discharge diagnosis of CoA at LUH in Reykjavik between 1984 and 2015. Based on evaluation of patient records we excluded two individuals who were non-Icelanders and two individuals who received an incorrect discharge diagnosis of CoA. Thus, the CoA sample set included a total of 132 CoA cases.

Information from the Book of Icelanders on gender and date of birth was available for 131 cases. There were 81 males and 50 females (Table 1), yielding a male:female ratio of 1.62:1. The cases received an initial diagnosis of CoA between 1950 and 2015 (Figure 2). There are very few diagnoses between 1950 and 1990 compared to 1990 and beyond, where 10 to 25 cases of CoA have been diagnosed per year at LUH.

Through patient’s records at LUH, we gathered detailed phenotypic information on all CoA cases and are the main phenotypes listed in Table 1. We excluded late hypertension from the study because accurate information on it was not available at LUH for majority of cases. This is because monitoring of late hypertension is usually performed at outpatient clinics. About half of the initial diagnoses were during the first month of life and about three quarters during the first year of life (Table 1). Individually, cases with moderate CoA were the most common. However, the combined occurrence of mild and moderate CoA is less than that of the more severe and complex types of CoA (Table 1). BAV is the most common associated CHD, occurring in 53% of cases. VSD is the second most common CHD, occurring in 35% of cases. Extracardiac birth defects, recoarctation and TH occurred in 24%, 17% and 61% of cases, respectively. Because of the young age of the sample set (data not shown), the rate of late cardiovascular complications was low, as exemplified with coronary artery disease (Table 1).
3.2 Genome-wide association analysis

To search for sequence variants that affect the risk of CoA GWAS was performed. The sequence variants (SNPs and INDELs) included in the GWAS analysis were identified through WGS of 8,453 Icelanders. The variants were subsequently imputed into 150,656 Icelanders, assisted by long-range phased haplotypes, who had been genotyped with Illumina SNP arrays. In addition, we used genealogical information to calculate genotype probabilities for 294,212 first and second-degree relatives of array genotyped individuals. Among the imputed Icelanders were 132 CoA cases identified as outlined above (see 3.1). GWAS case-control analysis was performed using the 132 CoA cases and 339,213 Icelanders with imputed genotypes that were used as controls. Association testing in the case-control analyses was performed using logistic regression, adjusting for gender, age and county. A total of 25.5 million variants were used in the association analysis. The threshold for genome-wide significance association was corrected for multiple testing using a class-specific Bonferroni procedure based on functional impact of classes of variants (74) and as outlined Methods (see 2.3). Using this set

| Table 1. Frequencies of various phenotypes among coarctation of the aorta cases |
|-----------------|---|---|---|
| Phenotype                  | N<sup>a</sup> | N<sup>b</sup> | Freq. (%)<sup>c</sup> |
| Male gender               | 131 | 81  | 61.8 |
| Diagnosed <1 month old    | 131 | 61  | 46.6 |
| Diagnosed 1 month to 1 year old | 131 | 33  | 25.2 |
| Diagnosed >1 year old     | 131 | 37  | 28.2 |
| Family history of CHD     | 90  | 24  | 26.7 |
| Extracardiac birth defect | 130 | 31  | 23.8 |
| Recoaartation             | 123 | 21  | 17.1 |
| Tubular hypoplasia<sup>d</sup> | 123 | 75 | 61.0 |
| Bicuspid aortic valve     | 124 | 66  | 53.2 |
| Ventricular septal defect<sup>e</sup> | 131 | 46 | 35.1 |
| Other CHD<sup>f</sup>      | 131 | 48  | 36.6 |
| Coronary artery disease   | 96  | 4   | 4.2  |
| Mild CoA                   | 132 | 11  | 8.3  |
| Moderate CoA               | 132 | 51  | 38.6 |
| Severe CoA                 | 132 | 34  | 25.8 |
| CoA with ventricular septal defect | 132 | 13 | 9.8 |
| CoA with complex CHD       | 132 | 19  | 14.4 |
| Atypical CoA               | 132 | 4   | 3.0  |

<sup>a</sup>N: the number of CoA cases with available phenotypic information for the particular phenotype.  
<sup>b</sup>N: number of CoA cases with particular phenotype.  
<sup>c</sup>Freq (%): the frequency of the particular phenotype among CoA cases.  
<sup>d</sup>Tubular hypoplasia may be mild to severe.  
<sup>e</sup>Ventricular septal defect may be tiny to large.  
<sup>f</sup>Other CHD includes all congenital cardiovascular defects except patent ductus arteriosus, persistent left superior vena cava and cardiomyopathy.
of thresholds, we observed a genome-wide significant association with CoA and variants located on chromosome 14q11 (Figure 3). No other variants in the genome reached the threshold of genome-wide significance (Figure 3).

**Figure 3. Manhattan plot of coarctation of the aorta genome-wide association study.** The $P$ values (-log10) are plotted against their respective positions on each chromosome. $P = 5 \times 10^{-8}$ is indicated by the horizontal dotted line. The arrow points to the c.2161C>T variant in MYH6. The plot was created using qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots.

**Figure 4. Region plot for the association of variants with coarctation of the aorta.** Shown is a 1Mb region on chromosome 14. The strongest association is with the missense variant c.2161C>T in MYH6 located at position 23,396,970 on chromosome 14. Nine other variants shown are weakly correlated with c.2161C>T, $r^2$ between 0.6-0.4 (green) and 0.4-0.2 (blue).
The strongest association with CoA at 14q11 is with a rare (frequency = 0.35%) missense variant c.2161C>T in MYH6, a gene encoding the alpha cardiac myosin heavy chain subunit (αMHC) protein. This variant associates under the multiplicative model with an odds ratio (OR) of 31.4 (95% confidence interval (CI); 14.08, 69.83) and with a P of 3.3 X 10^-17 (Figure 4). No homozygous carriers were found among CoA cases or in the general population (N = 150,656). This was expected because the variant is rare (frequency in the population 0.35%). Since no homozygotes were found we could not discriminate between dominant or multiplicative model of inheritance.

The c.2161C>T variant is located in exon 18 of MYH6 and leads to a missense arginine to tryptophan alteration at amino acid 721, p.Arg721Trp (Figure 5). The p.Arg721Trp mutation is in the converter domain of αMHC (Figure 5 and Figure 6), a small domain crucial in conveying a conformational change from the active site to the lever arm upon adenosine triphosphate (ATP) hydrolysis (76). Based on PolyPhen-2 (77), the p.Arg721Trp alteration is predicted to alter the structure of the converter.

![Figure 5. Transcript and protein structures of MYH6.](image)

![Figure 6. A 3D structural model of chicken smooth muscle myosin that is homologous to MHCα.](image)
The c.2161C>T mutation has previously been reported by deCODE to associate strongly with sick sinus syndrome (SSS) (66), a common disorder that is characterized by pathological sinus bradycardia (slow heart rate), sinus arrest and/or chronotropic incompetence (attenuated heart rate response to exercise) (78). The mutation also associated with pacemaker implantation (PM) but SSS is the most common indication for permanent pacemaker implantation (66).

The p.Arg721Trp mutation is neither present in The Exome Aggregation Consortium Browser (79), which holds sequence data from 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies, nor in the Exome Variant Server, holding sequence data from 13,000 individuals (80). The c.2161C>T mutation thus appears to be absent from other populations or if present, in a very low frequency. The Icelandic population is a founder population in that a small number of ancestors accounts for a relatively large proportion of the population. Hence, sequence variants that are very rare in more outbred populations may be much less so, like c.2161C>T, in Icelanders (68).

### 3.3 Phenotypic differences between carriers and non-carriers of c.2161C>T

To determine whether there are any phenotypic differences between non-carriers and carriers of the c.2161C>T mutation, we evaluated the frequencies of various phenotypes among the two groups of cases (Table 2). Of the 132 CoA cases, 24 were heterozygous carriers of the c.2161C>T mutation. As discussed above, no homozygous carriers were found among CoA cases or in the general population. Among CoA cases, information for the various phenotypes was available for 76 to 108 non-carriers and for 18 to 24 carriers (Table 2).

Because of the young age of the CoA sample set (3.1 above), we excluded late cardiovascular complications (few cases) such as coronary artery disease from the analysis. However, we included atrial septal defect (ASD), cardiomyopathy, and various arrhythmias as mutations in MYH6 have been associated with these diseases.

No significant phenotypic difference ($P<0.05$) was found between CoA non-carriers and carriers of c.2161C>T (Table 2). This may be due to constraints put on the analysis by the small size of the sample set. The small size reduced the power to detect significant difference between non-carriers and carriers and prohibited the individual analysis of some CHD that occurred with low frequencies. As an example, we did not analyze other LVOTOs individually (other than BAV).

Sick sinus syndrome (SSS) and heart block were the phenotypes with the largest frequency difference between non-carriers and carriers (Table 2). They were in elevated frequencies among carriers compared to non-carriers, with an OR of 4.08 ($P=0.36$) and 4.26 ($P=0.18$) for SSS and heart block, respectively (Table 2). The increased frequency, albeit not significant, of SSS among CoA cases that are carriers of c.2161C>T is in line with the reported association of c.2161C>T with SSS (66). In addition we observed that diagnosis after 1 year of age, mild or moderate CoA, pacemaker implantation (PM), cardiomyopathy, TH and recoarctation were about 2-times more common among CoA cases that are carriers of c.2161C>T compared to non-carriers (Table 2).
### 3.4 Association of c.2161C>T with other cardiac conditions

As outlined above, in a previous study by deCODE, the c.2161C>T mutation was found to associate strongly with SSS and PM. In that same study, residual associations, after exclusion of known cases of SSS, were reported with several cardiac diseases including atrial fibrillation (AF) and thoracic aortic aneurysm. From the time of that publication, which was in 2011, the phenotype and genotype database

#### Table 2. Frequencies of various phenotypes among non-carriers and carriers of c.2161C>T

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Non-carriers</th>
<th>Carriers</th>
<th>N&lt;br&gt;^a</th>
<th>Freq. (%)&lt;br&gt;^b</th>
<th>N&lt;br&gt;^a</th>
<th>Freq. (%)&lt;br&gt;^b</th>
<th>OR&lt;br&gt;^c</th>
<th>P&lt;br&gt;^d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gender</td>
<td>107</td>
<td>24</td>
<td>54.2</td>
<td>1.47</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosed &lt;1 month old</td>
<td>107</td>
<td>24</td>
<td>54.2</td>
<td>1.47</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosed 1 month to 1 year old</td>
<td>107</td>
<td>24</td>
<td>54.2</td>
<td>1.47</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosed &gt;1 year old</td>
<td>107</td>
<td>24</td>
<td>54.2</td>
<td>1.47</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family history of CHD</td>
<td>72</td>
<td>18</td>
<td>22.2</td>
<td>0.77</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extracardiac birth defect</td>
<td>106</td>
<td>24</td>
<td>25.0</td>
<td>1.08</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tubular hypoplasia&lt;sup&gt;e&lt;/sup&gt;</td>
<td>101</td>
<td>22</td>
<td>72.7</td>
<td>1.89</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bicuspid aortic valve</td>
<td>102</td>
<td>22</td>
<td>59.1</td>
<td>1.33</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventricular septal defect&lt;sup&gt;f&lt;/sup&gt;</td>
<td>107</td>
<td>24</td>
<td>33.3</td>
<td>0.91</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial septal defect&lt;sup&gt;g&lt;/sup&gt;</td>
<td>106</td>
<td>24</td>
<td>12.5</td>
<td>1.12</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other CHD&lt;sup&gt;h&lt;/sup&gt;</td>
<td>107</td>
<td>24</td>
<td>41.7</td>
<td>1.29</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>76</td>
<td>19</td>
<td>5.3</td>
<td>2.04</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>76</td>
<td>19</td>
<td>0.0</td>
<td>0</td>
<td>0.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventricular fibrillation</td>
<td>77</td>
<td>19</td>
<td>5.3</td>
<td>1.01</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sick sinus syndrome</td>
<td>76</td>
<td>19</td>
<td>5.3</td>
<td>4.08</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart block</td>
<td>76</td>
<td>19</td>
<td>10.5</td>
<td>4.26</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacemaker implantation</td>
<td>76</td>
<td>19</td>
<td>10.5</td>
<td>2.1</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild CoA</td>
<td>108</td>
<td>24</td>
<td>12.5</td>
<td>1.78</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate CoA</td>
<td>108</td>
<td>24</td>
<td>50.0</td>
<td>1.76</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe CoA</td>
<td>108</td>
<td>24</td>
<td>16.7</td>
<td>0.52</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoA with ventricular septal defect</td>
<td>108</td>
<td>24</td>
<td>4.2</td>
<td>0.35</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoA with complex CHD</td>
<td>108</td>
<td>24</td>
<td>12.5</td>
<td>0.82</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical CoA</td>
<td>108</td>
<td>24</td>
<td>4.2</td>
<td>1.52</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>N: the number of CoA cases with available information for the particular phenotype among non-carriers and carriers.  
<sup>b</sup>Freq (%): the frequency of the particular phenotype among non-carriers and carriers.  
<sup>c</sup>OR: Odds ratio.  
<sup>d</sup>P: Student's t-test was used to evaluate the significance of the frequency difference between non-carriers and carriers.  
<sup>e</sup>Tubular hypoplasia may be mild to severe.  
<sup>f</sup>Ventricular septal defect may be tiny to large.  
<sup>g</sup>Atrial septal defect does not include patent foramen ovale and tiny atrial septal defect.  
<sup>h</sup>Other CHD includes all congenital cardiovascular defects except patent ductus arteriosus, persistent left superior vena cava and cardiomyopathy.
at deCODE has expanded, with an increase in both the number of cases and controls for the various cardiac diseases. To gain a better understanding of the biological effect of the c.2161C>T mutation in MYH6 the association of the mutation with the various cardiac phenotypes in deCODE’s database was tested (Table 3). In this new dataset, the strongest associations are with SSS and PM, and both associations have become stronger when compared to the published association results ($P = 1.4 \times 10^{-49}$ vs $P = 1.5 \times 10^{-29}$ for SSS and $P = 23.6 \times 10^{-39}$ vs $P = 3.6 \times 10^{-25}$ for PM). In addition, the association with AF is now of genome-wide significance (Table 3).

Table 3. Association of c.2161C>T with various cardiac conditions

<table>
<thead>
<tr>
<th>Cardiac traits</th>
<th>Ncases./Ncontr</th>
<th>OR (95% CI)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sick sinus syndrome</td>
<td>2,396/284,109</td>
<td>8.5 (6.37, 11.22)</td>
<td>$1.4 \times 10^{-49}$</td>
</tr>
<tr>
<td>Pacemaker implantation</td>
<td>2,389/336,661</td>
<td>7.3 (5.39, 9.75)</td>
<td>$3.6 \times 10^{-39}$</td>
</tr>
<tr>
<td>Coarctation of the aorta</td>
<td>132/339,213</td>
<td>31.4 (14.08, 69.83)</td>
<td>$3.3 \times 10^{-17}$</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>10,718/329,896</td>
<td>2.5 (1.98, 3.22)</td>
<td>$7.6 \times 10^{-14}$</td>
</tr>
<tr>
<td>Congenital heart disease</td>
<td>2,465/331,808</td>
<td>2.3 (1.6, 3.38)</td>
<td>$3.2 \times 10^{-05}$</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>1,551/186,663</td>
<td>2.4 (1.5, 3.89)</td>
<td>0.00019</td>
</tr>
<tr>
<td>Aortic stenosis</td>
<td>1,668/272,839</td>
<td>2.4 (1.5, 3.92)</td>
<td>0.00042</td>
</tr>
<tr>
<td>Hypertrophic cardiomyopathy</td>
<td>158/231,600</td>
<td>1.0 (1.1, 1.83)</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Shown are the number of affected individuals and controls used in the association analysis for each of the traits. \(^b\) Estimated OR and 95% CI for the association with c.2161C>T.

Mutations in MYH6 have previously been linked to dilated (DCM) and hypertrophic cardiomyopathy (HCM) (81) along with various other CHD (82, 83), of which the strongest association is with familial ASD (43). In all instances these mutations have been restricted to sporadic cases or to one or few families. We do not observe association of c.2161C>T with HCM (Table 3) or ASD (data not shown) in Iceland. The association detected with ischemic stroke (Table 3) is likely secondary to AF, its main risk factor (84). In this study c.2161C>T associates suggestively with aortic stenosis (AS). AS can be congenital with or without BAV or it can develop late in life with or without BAV. The AS sample set used in this study consists mainly of AS patients that developed the disease late in life. We are lacking information on the occurrence of BAV in these samples.

Based on these data it can be concluded that the c.2161C>T mutation in MYH6 appears to mainly affect the risk of cardiac conduction diseases such as SSS, PM and AF on one hand and the CHD CoA on the other hand.
4. Conclusions and Discussion

4.1 Coarctation of the aorta in Iceland

CoA is a congenital narrowing of the aorta and accounts for 3.8% of all congenital heart disease in Iceland (5). The site of CoA is usually at or near the point of the insertion of the DA, just distal to the origin of the left subclavian artery. Despite excellent surgical outcomes, CoA can be a life-long disease with high rates of long-term cardiovascular complications.

As a tertiary referral hospital and the only hospital in Iceland with pediatric cardiologists, all CoA cases in Iceland are referred to LUH. We identified 132 Icelanders (population of about 333,000) who received a discharge diagnosis of CoA at LUH in Reykjavik between 1984 and 2015.

As expected, CoA was about two times more common in males than in females (51). There are very few diagnoses between 1950 and 1990 compared to 1990 and beyond, where 10 to 25 cases of CoA have been diagnosed per year at LUH. This may be explained by several factors. First, our study involved individuals who received a discharge diagnosis, either an initial diagnosis or during follow-up, between 1984 and 2015. Thus, for those who were diagnosed before 1984, only those who survived and underwent follow-up at LUH after 1984 were included. In contrast, for those who were diagnosed after 1984, an initial diagnosis of CoA was enough to be included. Second, major advancements in cardiovascular medicine and surgery in the last few decades have led to more diagnoses and improved survival. Third, in 1986, a leading pediatric cardiologist began practicing at LUH. In our opinion, this is probably the most important reason for the large increase in diagnoses around 1990. Before him, only two pediatric cardiologists had been at LUH for separate and short periods of time. Since 1997, there have always been a minimum of two pediatric cardiologists on staff at LUH. We speculate that neonatal mortality resulting from either diagnosed or undiagnosed severe CoA accounts for the majority of lost cases between 1950 and 1990.

About half of the initial diagnoses were during the first month of life and about three quarters during the first year of life. Individually, cases with moderate CoA were the most common. However, the combined occurrence of mild and moderate CoA is less than that of the more severe and complex types of CoA. BAV is the most common associated CHD, occurring in 53% of cases. BAV is well recognized as the most common CHD associated with CoA and, as in our study, occurs in about 50% of patients (9, 10). VSD is the second most common associated CHD. This was expected but our figure of 35% is higher than the commonly cited figure of 20%. This is because we included in our study the very common tiny muscular VSDs, of which are typically excluded from other studies (9, 10). The frequency of extracardiac birth defects in CoA cases is in line with the classically reported figure of 25% in other studies (12). The frequency of recoarctation is 17%. In most other studies the frequency of recoarctation ranges between 3 to 15% (85) but has been reported to be as high as 41% (86). The high occurrence of TH correlates with its common association with CoA.
4.2 MYH6 c.2161C>T explains a large fraction of coarctation of the aorta in Iceland

The underlying genetic basis and pathogenesis of CoA is largely unknown. Through a GWAS of CoA in Iceland we found a rare (frequency = 0.35%) missense variant c.2161C>T in the MYH6 gene that associates with the disease. The c.2161C>T associates with CoA with large effect, an OR of 31.4 and with high significance, \( P = 3.3 \times 10^{-17} \) (genome wide significance threshold for missense variants is \( 6.5 \times 10^{-8} \)). The MYH6 gene has not previously been reported to associate with CoA although other very rare mutations in MYH6 have been linked to both DCM and HCM (81) and various CHD (82, 83), with the strongest association with familial ASD (43). These mutations have been restricted to one proband or one to few families. In contrast, the c.2161C>T mutation explains a large fraction or 19% of CoA cases in Iceland and is thus an important candidate gene for CoA predisposition. No significant phenotypic differences were found between CoA carriers and non-carriers of c.2161C>T. This may in part be explained by the small size of the study. Although the c.2161C>T mutation could not be detected outside of Iceland other mutations within MYH6 may predispose to CoA in other populations. The high frequency of c.2161C>T in Iceland indicates that the mutation is likely a founder mutation.

4.3 Pathogenesis of MYH6 c.2161C>T in coarctation of the aorta

Cardiac muscle myosin is a cardiac structural protein and to our knowledge such a protein has not been linked to CoA before. Cardiac muscle myosin is a predominant component of the sarcomere, the basic contractile unit of cardiac muscle (66). It is a hexamer consisting of two heavy chain subunits, αMHC and beta myosin heavy chain (βMHC), two light chain subunits and two regulatory subunits. The αMHC heavy chain subunit is encoded by MYH6 whereas βMHC is encoded by MYH7. βMHC is the heavy chain isoform predominantly expressed in human heart whereas αMHC is highly expressed throughout life in the atrium and, additionally, alongside MYH7 in the ventricle during embryonic cardiogenesis (87). Expression of MYH6 has not been detected in the aorta (88). MYH7 is a well-established causative gene for cardiomyopathy (89, 90) and has been linked to several CHD (44, 91) whereas a more limited number of MYH6 mutations have been reported in DCM and HCM and in CHD, such as ASD as outlined above. In the myosin molecule, βMHC acts as a relatively slow ATPase whereas αMHC is a fast ATPase. In heart failure and other cardiac disorders in humans, βMHC is upregulated whereas αMHC is downregulated, resulting in reduction of cardiac performance (92) and it has been suggested that even minor shifts in the ratio between αMHC and βMHC subunits can markedly influence cardiac function.

The c.2161C>T mutation changes arginine to tryptophan at amino acid 721 (p.Arg721Trp) that lies in the converter domain of αMHC. This domain functions as a socket for the C-terminal α-helical tail of the αMHC and plays a critical role in amplifying the structural rearrangements in the motor domain and transmitting them to the α-helical tail during movements of the myosin during contraction (76). Based on PolyPhen-2 (77), the p.Arg721Trp alteration is predicted to alter the structure of the converter. However, whether this missense mutation increases or decreases activity of αMHC, both of which can alter sarcomere activity, is pending further structure-function analysis of the protein. In mice it has been shown that the cardiomyopathy mutation R403Q in MYH6 (93), increases sarcomere activity. In contrast, as outlined above, αMHC downregulation (decreased activity) has been linked to reduced cardiac
performance. The c.2161C>T mutation could thus predispose to CoA either through increased or decreased activity of the sarcomere.

As outlined in the Introduction (see 1.2), the exact mechanism by which CoA is produced is not clearly understood. Two theories have gained the widest recognition in explaining this, the hemodynamic theory and the ductal tissue theory. According to the hemodynamic theory, lesions that result in deceased left ventricular outflow promote development of CoA by reducing blood flow through the fetal aortic isthmus (29). In our opinion, the hemodynamic theory fits best with the pathogenesis of the c.2161C>T mutation. We propose a variation on the theory and suggest that the c.2161C>T mutation could predispose to CoA by reducing blood flow through the fetal aortic isthmus because of diminished contraction of the developing heart. Our hypothesis is compatible with the fact that MYH6 is expressed in the ventricle during embryonic cardiogenesis (87) and that expression has not been detected in the aorta or the DA (88). If the MYH6 c.2161C>T mutation were to predispose to CoA because of decreased left ventricular outflow it would be predicted that the mutation reduces rather than increases sarcomere activity.

4.4 c.2161C>T in other cardiac conditions

The large number of cardiac phenotypes and samples in deCODE’s database allowed us to perform phenotype-genotype correlation of the c.2161C>T mutation. Based on these analyses it is clear that in addition to CoA the mutation associates strongly with arrhythmias like SSS, PM and AF. The mutation does not appear to associate strongly with HCM or CHD apart from CoA.

The hemodynamic theory can be used to explain how the MYH6 mutation predisposes to CoA but how can we explain the association with arrhythmias that are diseases of the elderly? Patients who undergo successful repair of CoA remain at high risk for developing various late cardiovascular complications such as hypertension, premature coronary artery disease and cerebrovascular accidents. In contrast, arrhythmias have not been described as common cardiovascular complications in patients with CoA, suggesting that they do not develop secondary to CoA but rather are the result of altered activity of the sarcomere and contraction of the heart.

4.5 Future studies

In order to gain a better understanding of the pathogenesis of c.2161C>T in CoA, an important next step would be to perform structure-function studies of the c.2161C>T mutation. The induced pluripotent stem cell (iPSC) system is an interesting model to study the effect of mutations on cardiomyocyte contraction. In this system human skin fibroblasts from carriers of c.2161C>T would be engineered into iPSCs that could then be differentiated in vitro, under the correct circumstances, into cardiomyocytes (94). The contraction of cardiomyocytes with and without c.2161C>T could then be compared in vitro.

CoA is classified as a LVOTO and is thought to share genetics and developmental pathways with other LVOTOs. It would be very interesting as a next step to determine if that holds for the c.2161C>T mutation by performing individual GWAS for the other LVOTOs using the data set available at deCODE.
5. References


