Review

Studies of Genes Involved in Congenital Heart Disease

Tushar K. Ghosh 1, Javier T. Granados-Riveron 2, Sarah Buxton 1, Kerry Setchfield 1, Siobhan Loughna 1 and J. David Brook 1,*

1 Institute of Genetics, School of Life Sciences, Queen’s Medical Centre, University of Nottingham, Nottingham, NG7 2UH, UK; E-Mails: tushar.ghosh@nottingham.ac.uk (T.K.G.); plzsb2@exmail.nottingham.ac.uk (S.B.); plzks1@exmail.nottingham.ac.uk (K.S.); siobhan.loughna@nottingham.ac.uk (S.L.)

2 Laboratory of Genomics, Genetics and Bioinformatics, Hemato-Oncology and Research Unit, Hospital Infantil de México Federico Gómez, Dr Márquez 162, Col. Doctores 06720, Mexico City, Mexico; E-Mail: javier.granados@himfg.edu.mx

* Author to whom correspondence should be addressed; E-Mail: david.brook@nottingham.ac.uk; Tel.: +44-115-823-0345.

Received: 7 February 2014; in revised form: 4 May 2014 / Accepted: 12 May 2014 / Published: 22 May 2014

Abstract: Congenital heart disease (CHD) affects the intricate structure and function of the heart and is one of the leading causes of death in newborns. The genetic basis of CHD is beginning to emerge. Our laboratory has been engaged in identifying mutations in genes linked to CHD both in families and in sporadic cases. Over the last two decades, we have employed linkage analysis, targeted gene sequencing and genome wide association studies to identify genes involved in CHDs. Cardiac specific genes that encode transcription factors and sarcomeric proteins have been identified and linked to CHD. Functional analysis of the relevant mutant proteins has established the molecular mechanisms of CHDs in our studies.

Keywords: congenital heart disease (CHD); heart development; cardiac transcription factors; cardiac sarcomere genes
1. Introduction

Heart development is a complex process that requires precise interaction of signalling molecules, transcription factors, co-factors and structural proteins. Both inherited and non-inherited factors are involved in congenital heart disease (CHD) [1]. The incidence of CHD is about 4–6/1000 live births and the true prevalence could be 40/1000 if bicuspid aortic valve is included [2,3]. Despite the progress in diagnosis and intervention our understanding of the cause and mechanism of CHD remains limited. Over the past two decades, linkage analysis and candidate-gene screening have helped to identify the genetic causes in some cases of CHD. With the advent of whole exome/genome sequencing it is very likely that many more CHD causing genes will be identified, which will expedite our understanding the genetic causes of the disease. Most importantly, next generation sequence technology will enable us to study sporadic cases, the most common presentation of CHD. Targeted deletions in mice suggest there are more than 500 genes involved in heart disorders (http://www.informatics.jax.org/). Thus one can expect a similar number of genes taking part in heart development and disorders in humans. In this review we focused on the contributions of cardiac transcription factors and sarcomeric protein encoding genes in CHDs. The role of signaling and signaling-related molecules in CHD can be found elsewhere.

2. Transcription Factor TBX5

Cardiac transcription factors are the key regulators in heart development and mutations in the genes encoding these are a major cause of CHD [4,5]. Whilst studying the cause of Holt-Oram syndrome (HOS), an autosomal dominant disorder that predominantly affects the heart and the upper limbs, we identified mutations in TBX5 which encodes a T-box transcription factor as the primary cause of HOS [6]. Furthermore we identified its DNA binding elements and characterised its targets [7]. Functional analysis of the TBX5 mutations established the molecular mechanism of disease. Most of the mutations result in TBX5 proteins that are defective in DNA binding, transcriptional activity and/-or protein-protein interactions. Other related studies also revealed the role of TBX5 in CHDs [8,9]. To date more than 70 mutations in this gene have been identified from CHD patients.

TBX5 interacting factors such as NKX2.5 and GATA4 also play important roles in heart development. Mutations have been mapped in the genes encoding these proteins in CHD patients with overlapping clinical features [10,11]. Other TBX5 interacting proteins have been identified that also play vital roles in heart development. One such protein is MEF2C, a MADS box transcription factor. It associates with TBX5 and is required for early heart development [12]. A novel sequence variant that leads to A103V in MEF2C has been identified from a patient with cardiac outflow tract (OFT) defects. Overexpression of the A103V MEF2C variant in fish is found to perturb early cardiac development suggesting a role in CHD [13].

Mutations in other T-box transcription factor genes are also implicated in CHD. For example TBX20 is associated with defects in septation, valvulogenesis and cardiomyopathy [14]. It also interacts physically and genetically with other transcription factors such as TBX5, NKX2.5 and GATA4 [14]. Mutations in another T-box gene, TBX1, cause Di-George syndrome in which CHD is a major manifestation [15]. Mutations in other cardiac transcription factor genes such as GATA6 [16], NKX2.6 [17] CITED2 [18], IRX4 [19], ZIC3 [20], SALL4 [21], FOXH1 [22], FOXP1 [23], ZFPM2 [24] and TFAP2B [25] are also linked to CHD.
Functional interaction of the cardiac factors is vital for normal development and mutations in any of the interacting proteins that disrupt their physical association lead to CHD. TBX5 and its interacting partners GATA4 and MEF2C are key determinants of cardiomyocyte differentiation [26]. Overexpression of these three factors converts fibroblasts into cardiomyocyte-like cells. Addition of other factors such as HAND2, MYCOD, ZFPM2, ESRRG and MESP1 further enhance the efficiency of differentiation [27,28] (Figure 1). This is a significant breakthrough towards future regeneration therapy in the diseased heart.

**Figure 1.** A schematic diagram of the cardiogenic program. Core complex consisting of cardiac transcription factors TBX5(T), GATA4 (G) and MEF2C (M) is sufficient to reprogram fibroblasts to cardiomyocytes. Other accessory factors (HAND2, ZFPM2, MYCOD, MESP1 and ESRRG) enhance the transformation efficiency.

### 3. Myosin Heavy Chain Alpha (MYH6)

Sarcomeric proteins play important structural and functional roles in cardiomyocytes. Mutations in the genes encoding these proteins cause cardiomyopathy [29]. Myosin heavy chain alpha (MYH6) is one of the sarcomeric proteins implicated in cardiomyopathy [30,31]. Although not common, genes encoding structural proteins are emerging as new targets for studying CHD (Table 1). Using a genome-wide genetic linkage strategy employing microsatellite markers, we identified a region in chromosome 14q12 harboring a mutation causing secundum atrial septal defect in a large family in which the defect segregated as an autosomal dominant trait with incomplete penetrance. Sequence analysis of genes contained in that interval revealed a non-synonymous mutation in the gene encoding the α-cardiac myosin heavy chain (MYH6). This mutation consists on a single nucleotide substitution located in the segment encoding the neck domain of MYH6, which results in a I820N change in the mutant protein [32]. By surface plasmon resonance and *in vitro* pull-down assay, we showed that the mutant polypeptide has a decreased affinity for the atrial myosin regulatory light chain (MYL7) as its neck domain impose a stearic hindrance over the interaction between the two proteins [32]. This observation is consistent with studies in animal models showing that deficiency of the atrial myosin heavy chain resulted in altered cardiogenesis in zebrafish [33], *Xenopus* [34] and chick [35]. By mutational analysis using denaturing high-performance chromatography we screened a large cohort of
patients with apparently sporadic congenital heart defects for mutations in \textit{MYH6} [36]. This analysis allowed the identification of several private functional variants. Interestingly, two of them, A230P and A1366D decreased the capacity of MYH6 to form myofibrils, whereas H252Q has an enhanced potential to form these supra-molecular structures [36]. A later analysis by array-based resequencing identified \textit{MYH6} as the predominant sarcomeric disease gene for familial atrial septal defects [37] whereas by exonic sequencing of a large family with different types of congenital heart defect Arrington and collaborators also identified the A230P mutation [38]. Other mutations in \textit{MYH6} have been related to other human phenotypes like hypertrophic and dilated cardiomyopathy [39–43] and genomic variants within the gene have also been identified as susceptibility factors for complex traits such as heart rate [44,45] and sick sinus syndrome [46].

\textbf{Table 1.} List of genes encoding transcription factors and structural proteins that are associated with CHDs in human. Abbreviations: AA-Aortic Aneurysm; AS-Aortic Stenosis; ASD-Atrial Septal Defect; AVSD-Atioventricular septal defects; CD-Conduction Defects; CHD-Congenital heart disease; CAVC-Common Atroioventricular Canal; CAT-Common arterial Trunk; DORV-Double Outlet Right Ventricle; HLV-Hypoplastic Left Ventricle; LVH-Left Ventricular Hypertrophy; LVNC-Left Ventricular Noncompaction; MVD-Mitral Valve Disease; PA-Pulmonary Atresia; PDA-Patent Ductus Arteriosus; PFO-Persistence of Foramen Ovale; PS-Pulmonary Stenosis; SVAS-Subvalvular Aortic Stenosis; TOF-Tetralogy of Fallot; TGA-Transposition of Great Arteries; VSD-Ventricular Septal Defect.

<table>
<thead>
<tr>
<th>Gene</th>
<th>CHD type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textbf{Transcription factors}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBX5</td>
<td>ASD, VSD, CD (Holt-oram syndrome)</td>
<td>[6]</td>
</tr>
<tr>
<td>TBX1</td>
<td>ASD, VSD, TOF, PA (DiGeorge syndrome)</td>
<td>[15]</td>
</tr>
<tr>
<td>TBX20</td>
<td>ASD, VSD, PFO, MVD</td>
<td>[14]</td>
</tr>
<tr>
<td>GATA4</td>
<td>ASD, VSD, AVSD</td>
<td>[11]</td>
</tr>
<tr>
<td>GATA6</td>
<td>PTA, TOF, ASD</td>
<td>[16]</td>
</tr>
<tr>
<td>NKX2.5</td>
<td>ASD, VSD, SVAS, LVH</td>
<td>[10]</td>
</tr>
<tr>
<td>NKX2.6</td>
<td>CAT</td>
<td>[17]</td>
</tr>
<tr>
<td>MEF2C</td>
<td>OFT</td>
<td>[13]</td>
</tr>
<tr>
<td>CITED2</td>
<td>ASD, VSD, PS, TOF</td>
<td>[18]</td>
</tr>
<tr>
<td>IRX4</td>
<td>VSD</td>
<td>[19]</td>
</tr>
<tr>
<td>ZIC3</td>
<td>Heterotaxy associated with CHDs</td>
<td>[20]</td>
</tr>
<tr>
<td>SALL4</td>
<td>Okihiro syndrome/VSD</td>
<td>[21]</td>
</tr>
<tr>
<td>FOXH1</td>
<td>TOF, TGA, DORV, CAVC, TA</td>
<td>[22]</td>
</tr>
<tr>
<td>FOXP1</td>
<td>AVSD and HLV</td>
<td>[23]</td>
</tr>
<tr>
<td>ZFPM2/FOG2</td>
<td>TOF</td>
<td>[24]</td>
</tr>
<tr>
<td>TFAP2B</td>
<td>Char syndrome, PDA</td>
<td>[25]</td>
</tr>
<tr>
<td>\textbf{Sarcomeric protein}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYH6</td>
<td>ASD</td>
<td>[32]</td>
</tr>
<tr>
<td>ACTC1</td>
<td>ASD</td>
<td>[53]</td>
</tr>
<tr>
<td>MYH7</td>
<td>VSD, LVNC</td>
<td>[47]</td>
</tr>
<tr>
<td>MYH11</td>
<td>VSD, AA</td>
<td>[48]</td>
</tr>
<tr>
<td>ELN</td>
<td>SVAS, PS, AS</td>
<td>[49]</td>
</tr>
</tbody>
</table>
Two other myosin heavy chain proteins MYH7 [47] and MYH11 [48] are also linked to CHD. Mutations in Elastin (ELN), one of the two components of elastic fibers also cause a form of CHD called supravalvular aortic stenosis [49].

4. Cardiac Actin Alpha (ACTC1)

Alpha cardiac actin (ACTC1) is another contractile protein in heart muscle cells linked to cardiomyopathy [50,51]. Pathological mutations in ACTC1 cause aberrant protein folding and perturbed filament formation [52]. The same gene is also linked to atrial septal defects. A substitution mutation M123V in ACTC1 has been identified from affected individuals in two large families and the M123V protein has reduced affinity for myosin [53]. Subsequent screening of 408 sporadic cases of CHD identified an ASD patient with 17bp deletion mutation in this gene, which is predicted to generate a non-functional protein. Morpholino-mediated knockdown of ACTC1 in chick shows delayed looping and reduced atrial septa suggesting its role in heart development [53].

5. Multiple Mutations and Heterozygosity

Heterogeneity of clinical phenotype is a common feature in congenital heart disease. Affected individuals having the same mutation can often have variable phenotypes, which suggest involvement of other factors. Recent studies point towards the role of mutations in more than one gene contributing to clinical severity and hypertrophic cardiomyopathy is one of the best studied models for multiple gene mutations [54,55]. In addition to single-gene Mendelian CHD, mutations in MYH6 may have a role in the pathogenesis of the oligogenic variety of the disease. Recently, we identified a family in which a private mutation of MYH6 (A1443D) has no evident effect in the mother of the proband. Another private mutation of the gene encoding the transcription factor NKX2-5 (L122P) is present in the father of the patient, again, with no noticeable effect. Their son, with secundum atrial septal defect, inherited both variants, which appear to have complementary effects on cardiogenesis [56]. Identification of such multiple mutations suggests an additive effect over the pathogenesis in our cases.

The distribution of mutations identified in TBX5, MYH6 and ACTC1 are summarized in Figure 2. It is believed that the mechanism by which the mutations cause CHD is haploinsufficiency (e.g., nonsense/frameshift mutations or mutations in the regulatory region) or altered dosage/function of the protein (e.g., mutations in the regulatory elements/indel/missense mutations). The correct dosage of proteins with intact functions is critical for proper development to proceed and any deviation from that regime will be disruptive and lead to CHD. The observation that defective sarcomeric proteins expressed during cardiac development cause heart malformation is supported by studies targeting genes encoding several contractile proteins in animal models and, interestingly, the same phenomenon occurs after experimental mutation of genes encoding endothelial shear-stress sensors and transducers as well as cardiac muscle contraction and tension-sensing proteins [57]. We have proposed that early intra-cardiac blood flow is an epigenetic factor in heart morphogenesis and that its defective generation (i.e., contraction), sensing and transduction could induce morphologic alterations of the organ [58]. Pronounced phenotypic variability is common between affected members of families with Mendelian CHD in the presence of the same mutation. Given the large number of genes implicated in CHD, it can be hypothesized that the effect of a great number of modifier loci could account for this observation.
Figure 2. Diagrams representing the TBX5, MYH6 and ACTC1 proteins and the changes induced by mutations described in the genes encoding them. The changes in the amino acid residue sequences are annotated with colour-coded arrows according to the associated phenotypes. The domains recognized within these proteins are also annotated.
6. Concluding Remarks

Because CHD behaves mainly as a complex trait, the proportion of familial cases determined by the Mendelian segregation of single-gene mutations is small. Although the study of candidate genes has yielded valuable information about the biology of the disease, the resulting findings only account for a reduced fraction of the total burden of CHD. Even though the relevance of some missense and splicing mutations that segregate through these families is clear when functional deficit is established by molecular approaches, the interpretation of the role and size of the individual contribution of some mutations remains a challenge. Also, as a consequence of the aforementioned paucity of variants found by candidate gene studies, there are very few reports of possible instances of digenic or oligogenic inheritance and therefore the size of the sample required to evaluate these scenarios comprehensively is difficult to estimate. By increasing the number of genes analysed simultaneously by massively parallel exome or genome sequencing approaches could potentially overcome some of these difficulties.

To develop a therapeutic strategy for CHD patients, a comprehensive understanding of genes involved in building the heart will be of utmost importance. Defining the molecular program early during development will be useful in regeneration of the defective heart. Identifying the CHD causing genes and their interacting networks will further be invaluable to build a future therapeutic strategy and afford better patient care.

Acknowledgments

This work is supported by the British Heart Foundation.

Conflicts of Interest

The authors declare no conflict of interest.

References


© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).