Molecular Genetics of Congenital Heart Disease and Holt-Oram Syndrome

By

Javier Tadeo Granados MD

Institute of Genetics

University of Nottingham

Nottingham

United Kingdom

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

October 2005

ACKNOWLEDGEMENTS

First I would like to thank to my supervisor David Brook for the huge and life-changing opportunity to learn and work in his lab and his guidance and support, far beyond the call of duty. I would also like to thank all members of David's lab, especially to Liz Packham and Steve Cross for valuable training, Mark Pope and Jacqueline Eason for sharing a lot of the work related to this thesis and Thelma Robinson for important practical assistance.

I am of course in eternal debt with my parents for their impeccable moral example and support. Nothing that I have managed to achieve would have happened without their help.

What I owe to my wife Laura, can be hardly expressed with words. She has been a superb companion in this adventure, an inexhaustible source of encouragement and the perfect balance for my lack of interest in mundane (but much needed) pursuits. She will never stop to amaze me with her resilience.

I would like to thank the Mexican people, which through a scholarship from the Mexican Council for Science and Technology (CONACYT) made possible for me to come and work for my PhD.

Nottingham, Bonfire Night, 2006.

ABSTRACT

Heart development is a complex process which is regulated by molecular mechanisms still largely unknown. Disruptions in these processes cause congenital heart defect, that affects over 1 out of every 100 live births and is responsible for most antenatal losses. In the last few decades, several mutations have been shown to cause isolated as well as syndromic congenital heart defects and the genetic contribution to this pathology now is being recognized as important not only for the rare familial cases but also in regard to the much more complex multifactorial varieties of the disease.

The work summarized in this thesis was mainly an effort to clarify the role of mutations of a particular gene, *MYH6*, in congenital heart disease. Recently, this gene was identified as responsible for a Mendelian variety of atrial septal defect.

The other main subject of this thesis is the mutational analysis work done in order to identify a new gene, besides *TBX5* and *SALL4*, for Holt-Oram Syndrome, a developmental disorder characterized for the coexistence of congenital heart defects with upper limb abnormalities. Four candidate genes within the most likely chromosomal interval have been screened and excluded as responsible genes.

Chapter 1 INTRODUCTION

1.1 The heart

1.1.1 Evolutionary aspects

Through evolution, hearts arose as devices to pump fluids through the body circulation and have varied greatly according to the size, biology and environment of the different organisms.

The first recognizable pattern of cardiac evolution begins with the contractile tubular organs found in the urochordates, like the modern ascidians. In order to provide some evolutionary perspective of the heart, the morphology and development of the heart-like systems in this organism as well as in *Drosophila melanogaster* will be discussed.

1.1.1.1 The heart in the ascidians

These rudimentary organs do not contain chambers, valves or antero-posterior polarity and show the ability to reverse peristaltic direction. The vessel is formed by a single layer of myoepithelial cells with no endothelial lining and surrounded by a pericardial coelom (Simoes-Costa et al. 2005).

The development of the ascidian pump has been studied in *Ciona intestinalis*. Two blastomeres expressing mesoderm posterior (MESP) a helix-loop-helix transcription factor, divide during gastrulation giving origin to two paired bilateral populations of cells expressing orthologues of genes involved in vertebrate cardiogenesis like Nkx, Gata and Hand. These cardiac rudiments migrate ventroanteriorly fusing along the ventral midline and forming the pump field in a similar way as the vertebrate primary heart field is constructed (Figure 1.1A) (Davidson and Levine 2003).

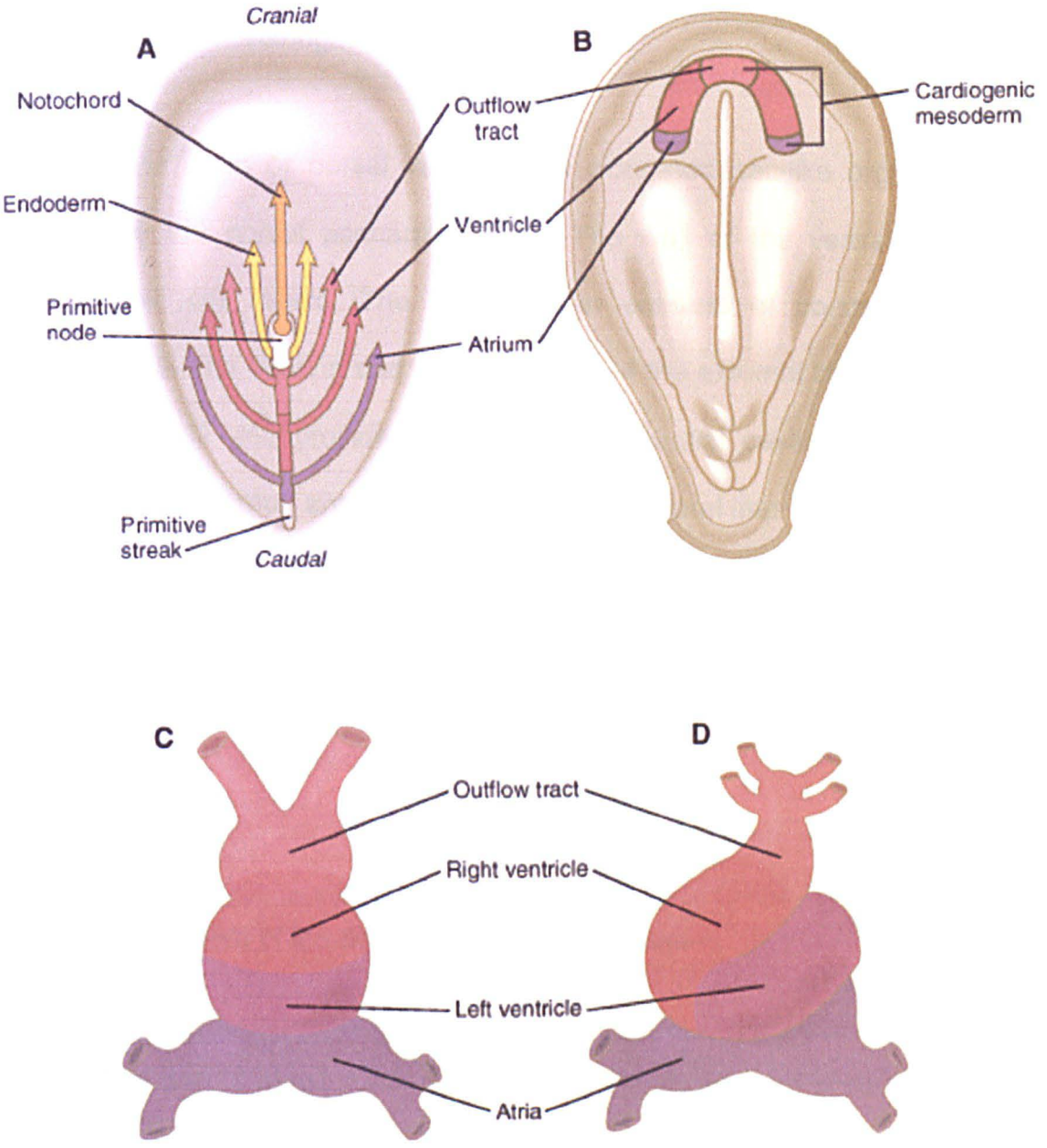


Figure 1.1: A, Localization and fate of the cardiac progenitor cells along the primitive line and their migration route to the cranial lateral plate mesoderm. B, Segmentation of the cardiac crescent according to the cardiac tissue they give rise to. C, The lineal cardiac tube before looping. D, The looped cardiac tube (see text). Figure taken from "Human embryology and developmental biology", by Bruce M. Carlson. 3rd ed. Mosby. Philadelphia (2004).

1.1.1.2 *Drosophila melanogaster*

The *Drosophila* heart or dorsal vessel is an elongated sac-like organ, located inside the dorsal pericardial cavity. The wall of the vessel is formed by two cell layers. The luminal layer consists of contractile myocardial cells. The filtering and secretory cells of the external layer are called nephrocytes and provide lining to the pericardial cavity. During relaxation the hemolymph enters the heart through several laterally placed ostia which are closed during contraction. This autonomous contraction wave is originated caudally and moves anteriorly. The hemolymph is expelled through the aorta which drains into the head of the animal (Bodmer 1999).

1.1.1.2.1 Development of the dorsal vessel in *Drosophila*

During the first part of stage 12 of development, the mesoderm tissue in the *Drosophila* embryo is divided in discrete masses of cells, the primordia of the somatic and visceral musculature, the fat body and the dorsal vessel. The cluster of cells that give rise to the dorsal vessel migrate dorsally beside the overlying ectoderm at stage 13, forming two rows of cells showing different features. The dorsal row cells form a straight line, assuming a cuboidal shape. These cells are the cardioblasts that form the contractile portions of the dorsal vessel. The cells of the lateral row

conserve their spherical shape and irregular arrangement, giving rise to the pericardial cells or nephrocytes (Campos-Ortega 1985). At stage 17 the row of cardioblasts of each side migrate centrally meeting and fusing to the contralateral. The resultant double row of cardioblasts forms a tubular structure with a central lumen.

Tinman (Tin), a *Drosophila* orthologue of the vertebrate family of homeobox transcription factors NK, has been shown to regulate the specification of the dorsal mesoderm (Azpiazu and Frasch 1993). This gene is initially expressed ubiquitously in the mesoderm cells under the control of the helix-loop-helix transcription factor *twist* (Yin et al. 1997). Its expression is limited subsequently to a section of the dorsal trunk mesoderm, maintained by induction of the overlaying endoderm by means of secretion of Decapentaplegic (Dpp) a member of the BMP family or transcription factors (Lockwood and Bodmer 2002). By stage 11 the expression of *tinman* is observed in dorsal clusters of mesodermal cells induced by neighbouring ectodermal and endodermal cells secreting *wingless*, a member of the Wnt protein family (Park et al. 1996). Both cardioblasts and nephrocytes derive from these cell clusters expressing *tinman* (Alvarez et al. 2003). In the absence of *tinman* expression or Dpp or Wnt signals, the dorsal vessel does not develop (Frasch 1995).

Pannier, a GATA transcription factor is a transcriptional target of Tinman in the heart forming region where they are coexpressed. Through physical interaction the two factors function synergistically in the regulation of

cardiac gene expression (Gajewski et al. 2001). During stage 11, *pannier* is induced by ectodermal Dpp signalling (Herranz and Morata 2001). The role of *pannier* in *Drosophila* heart development has been tested with mutants and ectopic expression experiments (Klinedinst and Bodmer 2003). Mesodermal *pannier* expression is in turn required to maintain adequate levels of *tinman* expression (Gajewski et al. 1999).

1.1.2 Heart development in vertebrates

Being designed for efficient and durable delivery of oxygen and nutrients to all tissues, the circulation must begin to function very early in development, and grow to meet the demands of the embryo.

As discussed above, heart development follows an evolutionarily conserved programme which is initiated by specific molecular signals and directed by tissue-specific transcription factors. This process controls the differentiation of mesodermal stem cells into cardiomyocytes and the subsequent expression of genes coding contractile proteins.

The complexity of the molecular determinants of heart development is highlighted by the sheer number of zebrafish mutants (Yelon 2001), null mouse mutants and human heritable conditions with altered heart morphogenesis (see below).

1.1.2.1 The cardiac progenitor cells

The first recognizable heart forming cells (around 50 in the mouse) (Tam et al. 1997) are found in the early gastrulating embryo as a linear cluster of epiblast cells along the anterior part of the primitive streak (Garcia-Martinez and Schoenwolf 1993). Derivatives of these cells contribute to form endocardium, myocardium and parietal pericardium. (Fishman and Chien 1997). The Cbp/p300-interacting transcriptional transactivator CITED2 (*MSG1*) has been shown to be a marker of these cells (Schlange et al. 2000a). Within this linear cluster, myocardial (Yatskievych et al. 1997) and heart segment specification already exist. The anterior-most cells will form the outflow tract while the posterior ones will contribute to the atria and inflow tract (Garcia-Martinez and Schoenwolf 1993). During gastrulation these cells, amongst others, invaginate through the primitive streak, between the epiblast and the definitive endoderm to form part of the lateral plate mesoderm (Rosenquist 1970) (Figure 1.1A). The transformation of the progenitor cells from the epiblast into mesoderm is thought to be initiated before gastrulation due to induction by the subjacent hypoblast through TGFβ superfamily molecules like Nodal or Vg1 (Cheng et al. 2003) and FGF signalling which seems to be important for the transition of their epithelial pattern in the epiblast to the mesenchymal one through repression on the cell-adhesion molecule E-cadherin via activation of the transcription factor *snail* (Ciruna and Rossant 2001).

The cardiac progenitor cells give origin to two different mesodermal cell lineages. The first constitutes the primary heart field whereas the second forms both the anterior and secondary heart fields (Meilhac et al. 2004).

Just after gastrulation, (15-16 days of development in humans, 7 dpc in mouse), the newly specified precardiac mesoderm cells of the first lineage migrate to form the cardiac crescent, a mesodermal structure in a horseshoe shape open caudally, located in the cranial regions of the embryo and containing diagonal bilateral and symmetric heart fields which joint medially and cranially (Schultheiss 1999). The migration route of these cells occurs according to their original position in the primitive streak. Those derived from the cranial part of the primitive streak migrate mainly cranially, close to the midline and occupy the anterior-most, fused pole of the crescent. Cells derived from the caudal end of the streak move cranially and laterally, in order to form two caudal ends of the crescent (Garcia-Martinez and Schoenwolf 1993) (Figure 1.1B).

This migration event seems to require the expression of *Fgf8* (Sun et al. 1999) and the basic helix-loop-helix proteins mesoderm posterior 1 and 2 (*Mesp1*, and *Mesp2*) (Kitajima et al. 2000).

1.1.2.2 The primary heart fields and the cardiac crescent

During the migration of the cardiac forming cells to form the cardiac crescent, they are contained in the local lateral plate mesoderm on the left

and right side and constitute the primary heart fields. The contribution of the heart fields to subsequent antero-posterior structures is unequal. The right heart field has a greater contribution to the cranial end of the heart tube whereas the left field contributes more to the caudal end (Stalsberg 1969) (Figure 1.1B).

When the lateral plate mesoderm divides in two epithelial layers, a dorsal one, called the somatic mesoderm and a ventral one, the splacnic mesoderm, the heart progenitors are segregated to the second (Linask et al. 1997) and express the calcium-dependent adhesion molecule N-cadherin (Takeichi 1991). The space left between the two mesodermal sheets constitutes the intra embryonic coelom which at the level of the cardiac crescent is called pericardial coelom, primordium of the pericardial cavity (Figure 1.2A). A number of vesicles between myocardium primordium cells begin to appear in the near vicinity of the pharyngeal endoderm. These vesicles eventually fuse to form the lumen of the two paired endocardial tubes and the N-cadherin negative cells lining the tubes constitute the endocardial primordium (Larsen 2001). The developing myocardium secretes an acellular and viscous matrix which is deposited between the myocardial and endocardial primordia, called the cardiac jelly (Carlson 2004) (Figure 1.2B).

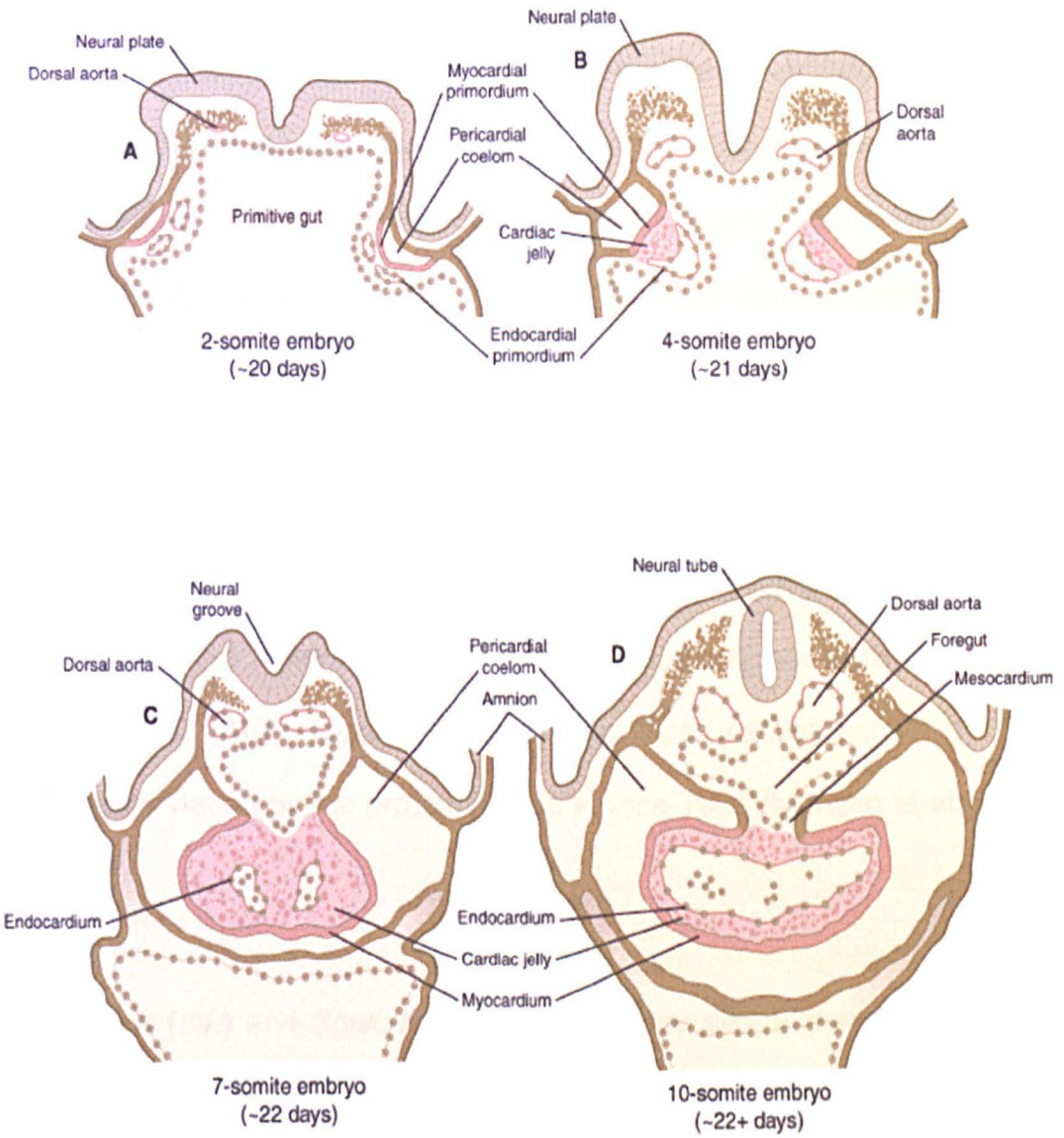


Figure 1.2: Diagram of transversal sections at four stages of development to show the structure of the heart forming structures and surrounding structures (see text). Figure taken from "Human embryology and developmental biology", by Bruce M. Carlson. 3rd ed. Mosby. Philadelphia (2004).

These morphological events in the cardiac crescent appear to be controlled by a number of inductions produced by the surrounding structures. The adjacent foregut endoderm secretes bone morphogenetic proteins (BMPs) producing an activity gradient increasing from the midline to the lateral edges of the crescent probably shaped by secretion of BMP inhibitors like Noggin and Chordin by the neighbouring axial mesoderm (Schlange et al. 2000b). The role of the BMPs in the crescent is to specify and maintain the myocardial lineage (Schultheiss et al. 1997) through positive regulation of the expression of Fgf8, required for cardiac induction (Alsan and Schultheiss 2002), and the Smad family (Liberatore et al. 2002). Smads in turn, interact with the basic-leucine-zipper activating trancription factor 2 (Atf2) (Monzen et al. 1999), which can be activated through phosphorylation by the protein kinase kinase Tak1 (Monzen et al. 2001).

Indian hedgehog (*Ihh*) and *Sonic hedgehog* (*Shh*) are also expressed by the pharyngeal endoderm below the crescent where their products along their receptor *smoothened* seem to function as permissive factors in cardiogenesis whereas their inhibitor receptor patched (Ptc) appears to have the opposite effect (Zhang et al. 2001).

Another kind of signalling molecules secreted by the endoderm adjacent to the myocardium primordium are Wnt inhibitors of the Dickkopf and Frizzled families (Schneider and Mercola 2001). Specifically, Crescent, one of the latter, binds to Wnt-1, Wnt-3, Wnt-8. All three molecules are present in the

neural ectoderm and expressed in the lateral plate mesoderm, acting through the canonical Wnt (via β-catenin) pathway, which is thought to negatively regulate specification of cardiac mesoderm (Marvin et al. 2001). Conversely, Wnt molecules acting through the non-canonical (PKC-dependent) pathway like Wnt-11, expressed in the posterior edge of the heart field, are powerful cardiac inducers (Koyanagi et al. 2005). The establishment of the lateral limits of the crescent is thought to be mediated by Notch signalling. The expression of one of its ligands, *Serrate,* is repressed in cardiogenic cells while differentiating, whereas is activated in the surrounding, non-cardiogenic mesoderm (Rones et al. 2000).

In response to the inductive stimuli produced by the neighbouring endoderm (BMPs, Fgf factors, Wnt inhibitors like Dkk and Crescent and Hedghog proteins), peripheral (Notch and Serrate) and central inhibitory influences (BMPs antagonist like Chordin and Noggin) by the notochord and posterior by the neural plate (β-catenin pathway Wnt proteins), the cardiac crescent activates a number of myocardial transcriptional regulators in the myocardium primordium, like Nkx2.5, Gata 4-6, myocyte transcription factors (Mef2b, Mef2c), heart and neural crest derivatives expressed 1 and 2 (Hand1, Hand2), T-Box 5 and 20 (Tbx5, Tbx20), serum response factor (SRF), Myocardin, besides genes encoding patterning molecules like the slow twitch cardiac muscle calcium transport ATPase (*Atp2a2*), phospholamban (*Pln*) and retinaldehyde dehydrogenase type 2 (*RALDH2*) (Brand 2003; Harvey 2002).

The expression of Gata6 on the crescent is restricted to the cells that are furthermost from the midline which, later in development, give origin to the heart conduction system (Davis et al. 2001).

Nkx2.5 is one of the first genes to be expressed in the cardiac fields, specifically in the cells with ventricular fate (Redkar et al. 2001) and it continues to be expressed in cardiac tissue throughout life (Komuro and Izumo 1993). Nkx2.5 is known to modulate the expression of other cardiac specific genes like *eHand* (Biben 1997a), Mef2c, N-myc, Msx2, myosin light chain 2V (Mlc2v) (Tanaka et al. 1999), atrial natriuretic factor precursor (*Nppa*) (Durocher et al. 1996), α-cardiac alpha (Chen et al. 1996), cardiac ankyrin repeat protein (*Carp*) (Zou et al. 1997), Iroquois homeobox gene 4 (*Irx4*) (Bruneau et al. 2000) endothelin-converting enzyme-1 (*ECE1*) (Funke-Kaiser et al. 2003), connexin-40 (*Cx40*) (Heathcote et al. 2005), type II iodothyronine deiodinase (DIO2) (Dentice et al. 2003), GATA4, TBX5 (Riazi et al. 2005; Sun et al. 2004), pro-collagen type I alpha 2 (pro-col1a1) (Ponticos et al. 2004), a cardiac-specific isoform of the RNA helicase Mov10l1 (Csm) (Ueyama et al. 2003). The list is growing fast and much more downstream targets of NKX2.5 are likely to be found in the future.

1.1.2.3 The anterior and secondary heart fields

The cardiac precursor cells from the second lineage are first identified as a second, Ffg8 and Fgf10 expressing crescent, continuous with and

medially located to the crescent of the primary heart field (Kelly et al. 2001). These cells give rise to a population of mesodermal cells located in the pharyngeal mesoderm that can be distinguished by their expression of the LIM homeodomain transcription factor Islet-1 (Cai et al. 2003), as well as Nkx2.5, Gata4 and Mef2c (Dodou et al. 2004; Waldo et al. 2001).

The anterior heart field comprises a subpopulation of pharyngeal mesodermal cells of the Islet-1 positive linage that express the forkhead transcription factor Foxh1 (von Both et al. 2004) and migrate through the splachnic mesoderm and the dorsal mesocardium in order to populate the prospective proximal outflow tract and the right ventricle during cardiac looping (Kelly et al. 2001; Mjaatvedt et al. 2001; Zaffran et al. 2004).

The secondary heart field is formed by another subpopulation of precardiac pharyngeal mesoderm cells that express Tbx1 and migrate to the heart tube to contribute to the formation of the distal outflow tract (Waldo et al. 2001; Xu et al. 2004) as well as the atria and the inflow tract (Brown et al. 2004).

1.1.2.4 The heart tube

Around the 22nd day of human development (8 dpc in the mouse), both cardiac primordia, each one formed by a medial endocardial tube, flanked laterally by a layer of cardiac jelly, myocardium primordium and the pericardial coelom are brought together in the midline driven by the lateral

and ventral folding of the embryo and fuse, creating a single, central heart tube (Carlson 2004) (Figure 1.1C and 1.2D).

This migration event depends on a cranio-caudal concentration gradient of fibronectin in the extracellular matrix between mesoderm and endoderm (Linask and Lash 1986) and a conserved structural and functional integrity of the foregut endoderm. The failure in central migration and fusion by the cardiac primordia produces two separate developing hearts, a condition called cardia bifida (Harvey 2002). The zebrafish mutants *casanova* (*sox32*) (Alexander et al. 1999), *faust* (*GATA5*) (Reiter et al. 1999) and the G-protein-coupled receptor for sphingosine-1-phosphate *miles apart* (Kupperman et al. 2000) and the mutant mouse embryos for *GATA4* (Narita et al. 1997), *Foxp4* (Li et al. 2004), *Mesp1* (Saga et al. 1999) and *furin* (Roebroek et al. 1998) display defects of the foregut endoderm and cardia bifida. Interestingly, *Foxp4* mutants develop two complete, separate hearts with proper looping and chamber formation, showing that fusion of the bilateral cardiac primordia is not required for cell or chamber specification in the heart (Li et al. 2004).

Outside the endocardial lining, the next layer of the heart tube is constituted by the cardiac jelly, followed by a layer of myocardial cells which remains attached dorsally to the pericardial mesoderm covering the foregut through the dorsal mesocardium, important as a migration route for the cells of the anterior and secondary heart fields (see above). The tube

is suspended inside the pericardial coelom by its connection to the arterial and venous vessels and the dorsal mesocardium (Figure 1.3).

The endocardial tubes do not fuse at their caudal end but continue posteriorly as the venous inflow tract of the heart whereas their cranial end generates arterial arches that surround the pharynx (Carlson 2004).

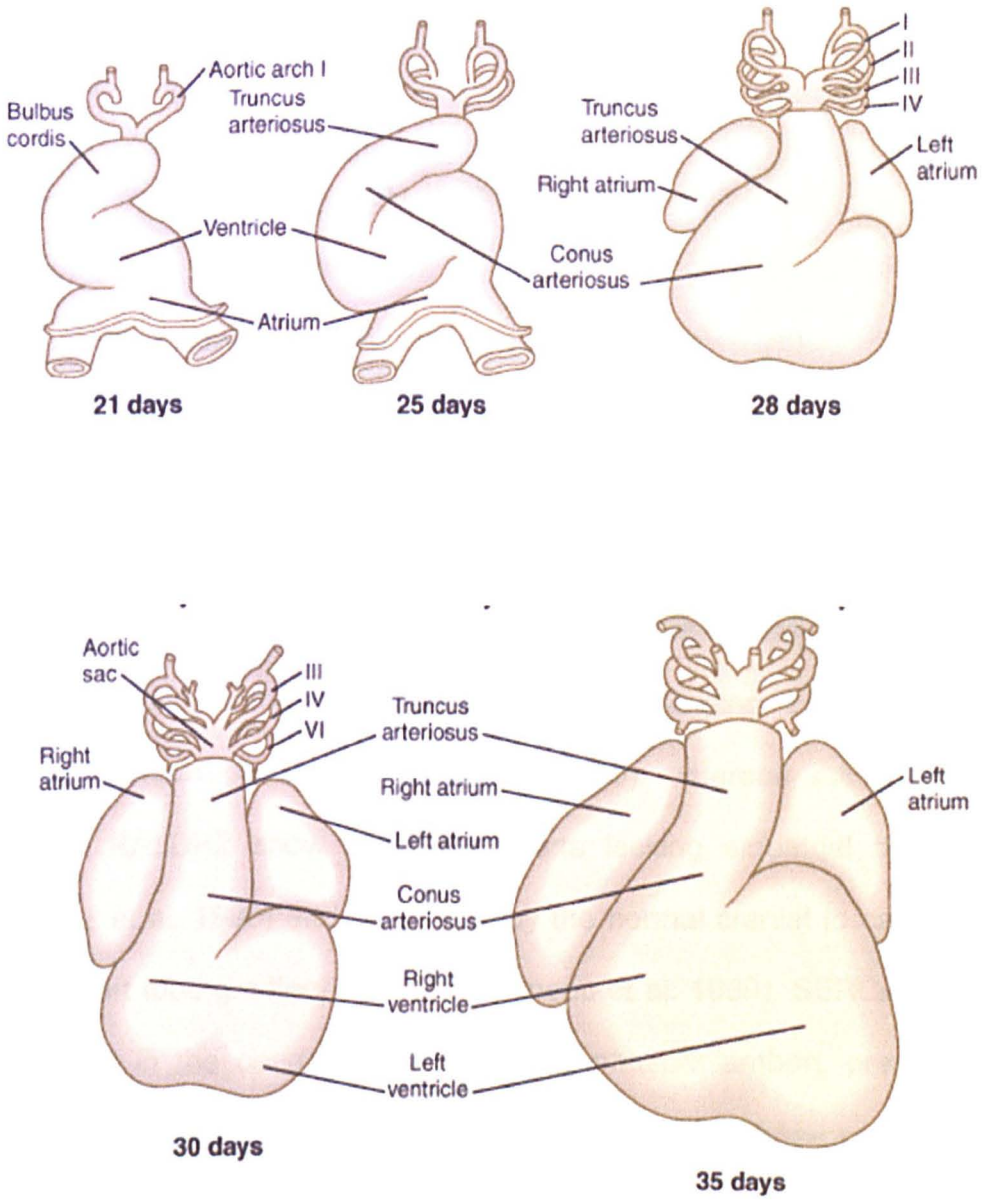


Figure 1.3: Diagram of transversal sections at four stages of development to show the heart and outflow tract (see text). Figure taken from "Human embryology and developmental biology", by Bruce M. Carlson. 3rd ed. Mosby. Philadelphia (2004).

1.1.2.4.1 Cranio-caudal patterning in the cardiac tube

Virtually all the retinoic acid (RA) in the embryo is synthesized by the retinaldehyde dehydrogenase type 2 (RALDH2). Expression of the *RALDH2* gene is limited to the sinuatrial regions during the cardiac crescent and tube stages. The resultant graded concentration of RA is thought to be important in the establishment of the cranio-caudal patterning seen in these structures. Indeed, high concentrations of exogenous RA in the embryo are deleterious to heart development only if the exposure occurs in the cardiac crescent stage, causing "atrialization" of the whole organ (Simoes-Costa et al. 2005), whereas mice with mutations in RALDH2 show unlooped hearts lacking sinuatrial tissue (Niederreither et al. 1996) and fail to display the normal cranial to caudal increasing heart tube gradient of TBX5 (Bruneau et al. 1999). SERCA2 is present mainly in the caudal inflow end, and phospholamban, present mainly in the cranial, outflow end (Moorman et al. 2000). These proteins regulate cytoplasmic calcium concentration and seem to be responsible for the highest rate of contraction in the caudal part of the tube.

1.1.2.5 Cardiac looping

The embryo displays a symmetrical configuration until gastrulation, when the symmetry is broken by the unidirectional movement of cilia from cells

around the primitive node (Nonaka et al. 1998). These posteriorly tilted cilia that move in clockwise direction (Nonaka et al. 2005) are thought to generate a leftward flow of yolk sac fluid (nodal flow) driving yet unidentified molecules to the left part of the embryo, where they could be responsible for the predominantly left lateral plate mesoderm expression of Nodal at the early somite stage, which in turn activates the expression of Lefty1 (along the left side of the primitive streak), Lefty2, and the bicoid-type homeobox Pitx2 in the same side (Nonaka et al. 2002), initiating a cascade of events that eventually produces the asymmetric morphogenesis seen in the adult organism. The nodal flow has been reversed experimentally in cultured mouse embryos. This resulted in a situs reversal with altered expression of several left-right determinants (Nonaka et al. 2002).

Heart looping represents the first macroscopically visual manifestation of left-right asymmetry in the vertebrate embryo. The heart tube develops in the closed pericardial cavity and in order to grow in length as well as width, it undergoes, between the 23rd and 28th day in humans, a rightward looping that ultimately places the primordia of the cardiac chambers in their definitive positions. The dorsal mesocardium, which initially tethers the developing cardiac tube, is disrupted, releasing most of it. As the looping progresses, the bulbus cordis (outflow tract, primordium of the conotruncus) is displaced caudally, ventrally and to the right. The prospective ventricle is displaced to the left and the primordium of the atria is displaced cranially and posteriorly (Figure 1.1D and 1.3). The ventral

surface of the tube is placed as the outer surface of the loop while the dorsal surface of the tube becomes the inner surface of the loop. When the heart chambers form, they begin to develop as evaginations of the outer surface of the looped heart, undergoing a process called ballooning.

The molecular mechanisms responsible of cardiac looping are not well understood. As embryonic hearts explanted in culture retain the ability to loop (Manning and McLachlan 1990), this phenomenon can not be explained solely by a constrained longitudinal growth. Marked differences in mitosis or apoptosis between regions of the looping heart tube have not been observed (Stalsberg 1970) and disruption of actin filaments (Itasaki et al. 1991) or microtubules (Icardo and Ojeda 1984) which directs cell shape do not prevent cardiac looping.

Cardiac looping appears to be controlled at least in part by the same pathway controlling laterality in other structures of the early embryo. The axial signalling system that determines the direction of cardiac looping also affects the position of the lungs, liver, spleen and gut. Interpretation of left-right signals is mediated in part by the transcription factor Ptx2, which is expressed along the left side of developing organs, including the early heart tube. Mouse models of left-right defects demonstrate absent, bilaterally symmetrical, or reversed Nodal and Ptx2 expression. In humans, mirror-image reversal of left-right asymmetry is often associated with normal organogenesis, but a discordance of cardiac, pulmonary and visceral asymmetry (heterotaxy syndrome) reflects a lack of coordinated

left-right signalling and is universally associated with defects in organogenesis. The common association of human cardiac alignment defects with abnormalities in left-right asymmetry points to intersecting pathways that regulate the direction and process of cardiac looping, and highlights the clinical significance of this area of study (Capdevila et al. 2000).

Very early in heart development, left-right differential gene expression can be detected. Transcriptional coactivator CITED2 expression is transiently stronger in the right heart field in the chick (Schlange et al. 2000a).

Other asymmetrically expressed genes during gastrulation and neurulation include activin receptor IIa (*ACVR2A*), *sonic hedgehog*, the the transcription factor *Foxa2* (Levin et al. 1995), *wnt-8c*, *patched*, the activin binding protein gene *follistatin* and the snail-related transcription factor gene *SnR-1* (Isaac et al. 1997; Levin 1997). Additionally, some extracellular matrix molecules like hLAMP and flectin (both expressed mainly on the left) and JB3 (expressed mainly on the right) are though to be important for looping as treatment with antibodies against flectin can randomize the direction of it (Linask et al. 2002).

Aberrant looping is observed in the *Nodal* hypomorph mouse (Lowe et al. 2001) and when this gene is ectopically expressed in the right side of the embryo (Levin et al. 1995).

During looping, expression of Hand1 occurs mainly on the left part of both atrial (Biben 1997a) and ventricular primordial and antisense translational inhibition of both Hand1 and Hand2 in chick embryos prevents normal looping (Srivastava et al. 1995).

In zebrafish, a Tbx20 antisense morpholino impairs looping, and a straight cardiac tube is formed (Szeto et al. 2002). In these morphants, expression of Tbx5 is upregulated whereas over expression of Tbx20 downregulates Tbx5. It is thought that these two proteins compete for the same binding sites on promoters of specific cardiac expressed genes like *ANF* (Plageman and Yutzey 2004). Moreover the zebrafish Tbx5 mutant *heartstrings* show unfolded hearts (Garrity et al. 2002).

In zebrafish cardiac looping is not the first macroscopic event breaking the symmetry of the early embryo. Instead, during cardiac "jogging" the caudal end of the heart tube is displaced in most embryos to the left side and then returns to the midline to start looping to the opposite side. Cardiac jogging always precedes looping and always occurs to the opposite side, even in mutants where looping is randomized. In a mutant screening in zebrafish, (Chen et al. 1997a), it was found that mutants showing altered looping morphogenesis display perturbations of the normally asymmetric *BMP4* expression, remaining symmetric or randomized. Those retaining *BMP4* symmetry show failure to "jog" whereas those with right-predominance of the *BMP4* pattern show reversion of the direction of jogging and looping. Experimental upregulation and downregulation of *BMP4* impairs

directional jogging or looping, suggesting that the expression pattern of *BMP4* could participate in the interpretation of laterality signals (Chen et al. 1997b).

1.1.2.6 Development of the venous inflow to the atria

By day 24$^{th.}$ the venous flow reaches the primitive atrium through two symmetric horns of the single and medial *sinus venosus*. Each horn in turn receives venous return through a common cardinal vein (draining the anterior and posterior cardinal veins), an umbilical vein (carrying oxygenated blood from the placenta) and a vitelline vein (draining the liver and other developing viscera). In the next days the venous system undergoes extensive reorganization. The three caudal tributaries of the left horn (posterior cardinal, umbilical and vitelline left veins) undergo involution, while the remains of the left horn and left common cardinal vein stop growing, and constitute later the coronary sinus and the oblique vein of the left atrium respectively. The right cardinal vein atrophies while the left umbilical and vitelline veins fuse to form the inferior vena cava and anterior cardinal vein gives rise to the superior vena cava (Larsen 2001).

The pulmonary vein sprouts from the primitive atrium at the beginning of the fourth week of development, as a strand of endothelium that runs through the dorsal mesocardium and makes connection with the vascular plexus surrounding the developing lung buds. The trunk of the pulmonary vein divides in left and right branches to each developing lung, which in turn bifurcate to produce a total of four pulmonary veins (Webb et al. 2001) (Figure 1.3).

1.1.2.7 Ballooning and specification of chamber myocardium

Anatomic identity of the cardiac cavities becomes to appear only soon after cardiac looping. However, the cell fate of the cardiac fibres of the chambers is genetically programmed since much earlier stages of development. The heart tube is divided along the anterior-posterior axis into precursors of the aortic sac, conotruncus pulmonary and systemic ventricles, and atria. Each one differs in its gene expression and functional patterns (Srivastava and Olson 2000).

The myocardial tissue that forms the primordia of the cardiac chambers is located after looping in the outer curvature of the cardiac tube. These segments experience a change called ballooning, which consists in the outpouching and thickening of the wall of the tube. Gene expression of this myocardial tissue as well as functional analysis suggest that this active or "working" myocardium generates most of the contractile effort and electrical connectivity of the organ at this stage.

A common feature of the more differentiated working myocardium is the development of trabeculae, which form the spongiform, irregular surface tissue lining the lumen of the ballooned segments of the cardiac loop, characteristic of the adult ventricles. Anf, Hand1, Cited1, Chisel, Irx1, Irx3, Irx5 Cx40 and Cx43 are expressed in the trabeculated areas of the outer curvature (Christoffels et al. 2000a; Christoffels et al. 2000b).

Between the endocardium layer and the prospective trabeculated layer of myocardium the cardiac jelly persists, though some areas of direct contact between the two layers exist. The development on the trabeculated tissue is induced by the subjacent endocardium by means of Neuregulin-1 molecules that diffuse through the jelly and reach the membrane tyrosine kinase co-receptors Erbb2 and Erbb4 of the myocardium (Lee et al. 1995). The functional interactions of these receptors with CD44 (Bourguignon et al. 2001), a membrane receptor for the glycosaminoglycans present in the jelly is thought to be important for the process, as trabeculae are absent in mice with mutations in the hialuronan synthetase-2 gene (Has2) (Camenisch et al. 2000).

Defects in ventricular trabeculation have also been observed in mice lacking angiogenic factors, such as vascular endothelial growth factor (VEGF) (Carmeliet et al. 1996) and angiopoietin-1 (Suri et al. 1996), that are expressed in the endocardium. In Nkx2.5 and Tbx5 null mouse

embryos the heart retains a tubular appearance, failing to loop and to express working myocardium markers (see above)

1.1.2.8 Development of the atria

After looping has occurred, the primordium of both atria is constituted by the cranially displaced caudal limb of the cardiac loop. The outer wall of this segment of the loop expands and thickens (in a process known as "ballooning") producing a single and medial chamber, the primitive atrium. This atrial component of the looped heart tube contributes in variable degree to the definitive atrial chambers (see below) and in both sites the cranio-lateral wall of the tube balloons to form both auricles, surrounding the developing outflow tract from either side (Moorman et al. 2003).

By the 31st day of development, the prospective superior and inferior venae cavae as well as the primordium of the coronary sinus empty into the right horn and the remaining left horn of the sinus venosus, which in turn drains into the primitive atrium through a long and narrow orifice located in the right posterior wall of the chamber (Carlson 2004).

The wall of the right horn of the sinus venosus is then incorporated to the right and posterior wall of the primitive atrium by an intussusception process, displacing the right part of the primitive atria wall ventrally and to the right and pulling the venae cavae and the coronary sinus, so each one drains at this stage through its own ostium. Two tissue flaps form at either

side of the three ostia, the left and right venous valves, and join cranially, to form the septum spurium. Later in development, the left valve is incorporated to the septum secundum, while the right valve forms the (Eustachian) valve of the inferior vena cava and the (Thebesian) valve of the coronary sinus (Larsen 2001).

The displaced part of the prospective right atrium forms the right or triangular auricle which has a broad connection with the atrium. The trabeculae in the auricle consist of pectinate muscles, which extend all around the parietal atrioventricular junction. The border between its trabeculated surface and the smooth wall of the rest of the chamber is indicated by the terminal crest which runs from its roof to the caudal atrioventricular junction and constitutes a preferential route for the electrical impulse between the sinuatrial and atrioventricular nodes. The left (tubular) auricle, which is formed by the same process in the opposite side, has a much narrower communication with the atrium and the trabeculated surface is limited to the auricle.

The differences in morphology between the two auricles are important because they can identify the morphologically right and left atria, regardless of their spatial arrangement in a malformed heart. This is because their primordia do not participate in cardiac looping and therefore they derive from separate left and right structures, in contrast to the ventricles, whose primordia are initially located in a cranio-caudal

sequence and receive contributions from both left and right elements (Campione et al. 2001).

Displaced myocardium from the primitive atrioventricular canal is an important source of atrial tissue. By the 10^{th} week of development, the caudal-most part of the atria, the smooth wall of the vestibule of both atrioventricular valves (Kim et al. 2001) is derived from remodelled myocardium displaced during the rightward movement of the atrioventricular canal.

Cells from the anterior and secondary heart fields that migrate during and after cardiac looping through the remaining dorsal mesocardium also contribute to the formation of the primary atrium (Meilhac et al. 2004).

During and after atrial septation the developing pulmonary vein that drains through a single ostium located in the cranial posterior wall of the prospective left atrium undergoes an intussuception process similar to that described for sinus venosus, incorporating the wall of the vessel to the roof of the chamber. The absorption reaches the first and second branching points of the pulmonary vein, so that ultimately four independent pulmonary veins communicate with the heart (Kelly et al. 2001).

These four sources of atrial tissue can be recognized as four different transcriptional domains. The auricles can be identified as a derivative of the pro-atrial segment of the cardiac tube because of their expression of

ANF in the fetal stages. MLC2V is expressed in the atrioventricular canal and sinus venosus derived tissues. Pitx2 specifically marks the interatrial septa, pulmonary vein derivatives as well as contributions of the left atrial auricle and left part of the atrioventricular canal (Franco et al. 2000) (Figure 1.4).

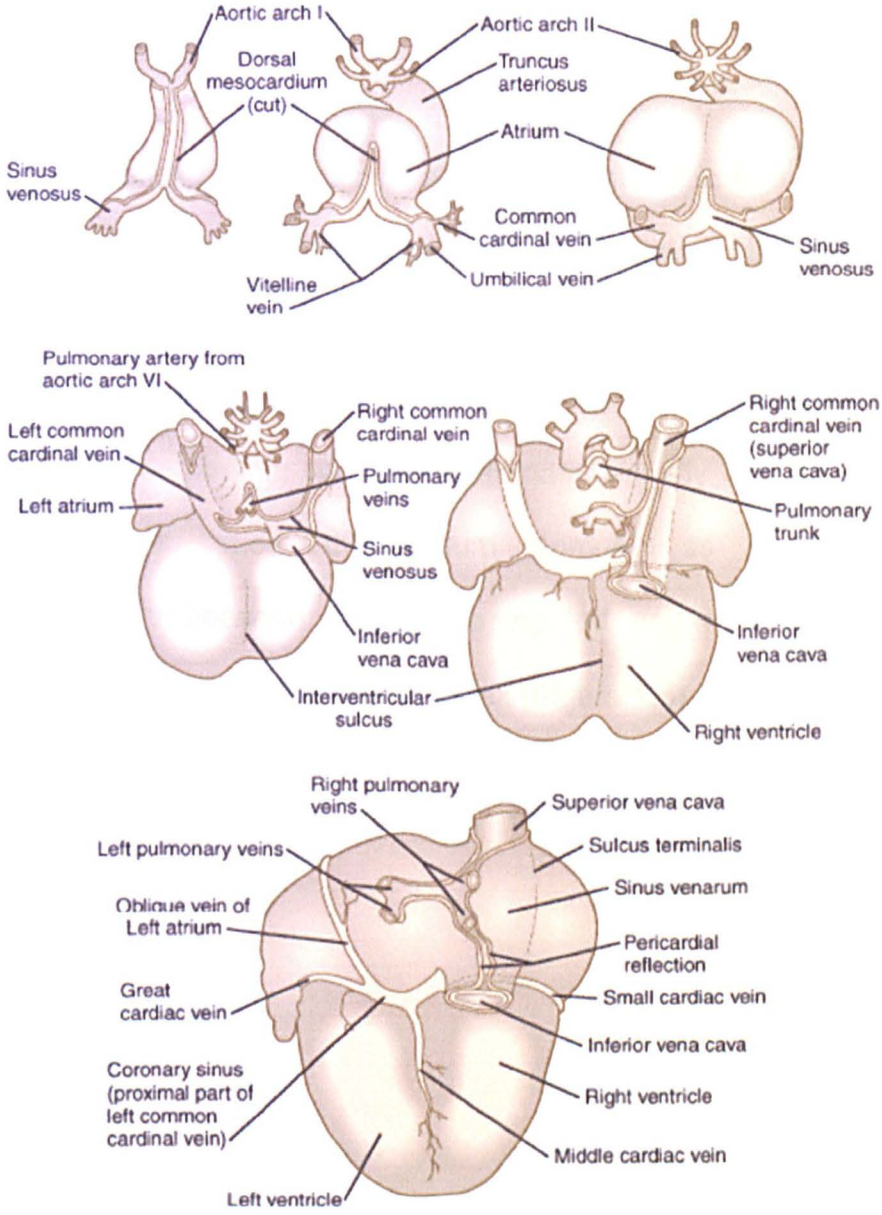


Figure 1.4: Diagram of the posterior aspect of the heart from just before looping to the four chamber stage showing the development of the venous inflow (see text). Figure taken from "Human embryology and developmental biology", by Bruce M. Carlson. 3rd ed. Mosby. Philadelphia (2004).

1.1.2.9 Septation of the atria

The division of the primitive atrium in a left and right atria is a complex process carried out in several stages approximately between the 33rd and 43rd days of development. Firstly a muscular septum develops in a crescent shape, emerging from the posterior and cranial walls of the primitive atrium. This laminar structure expands ventrally and caudally, leaving a rounded space, the ostium primum, as a transient communication point between the developing left and right atria that eventually closes near the end of the 6th week (Figure 1.5).

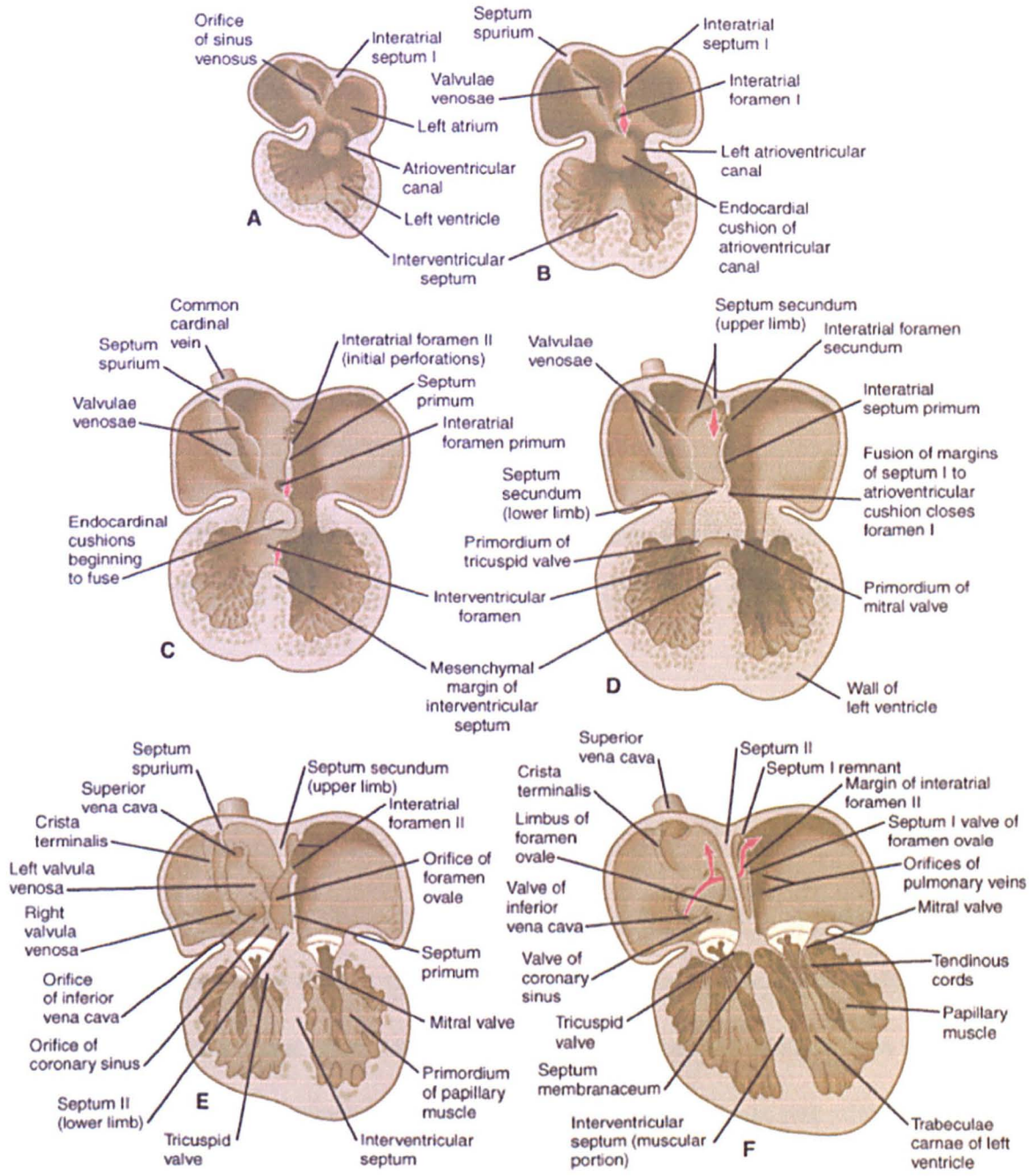


Figure 1.5: Diagram of coronal sections of the developing heart, showing the changes in the luminal morphology og the chambers, septa and valves (see text). Figure taken from "Human embryology and developmental biology", by Bruce M. Carlson. 3rd ed. Mosby. Philadelphia (2004).

The free edge of the expanding septum primum is covered by a cap formed by mesenchymal tissue, the cranial spur of the atrial spine. This relatively undifferentiated tissue has an extracardiac origin, entering the heart by the posterior mesocardium and penetrating the atrium to form the right rim of the pulmonary orifice and to be placed on top of the inferior endocardial cushion, producing its spur. When the spur and the body of the atrial spine fuse with the superior endocardial cushion, the ostium primum is obliterated. After the fusion, the atrial spine mesenchymal tissue becomes myocardium, except for the central portion, which condenses and persist as the fibrous tendon of Todaro (Webb et al. 1998).

The muscular tissue that form the septum primum expresses the left markers CK-B (Wessels et al. 2000) and Pitx2 (Campione et al. 2001).

Before the closure of the ostium primum is completed, a number of small perforations in the dorsal septum primum are formed by apoptosis, which eventually become confluent to constitute a bigger orifice, the ostium secundum, so the shunt between both developing atria is not interrupted.

While the left venous valve regresses, another infolding of the dorsal wall of the primitive atrium, the septum secundum, develops in a similar way as the septum primum, growing ventrally and caudally as a crescent and leaving a space that becomes a caudal and dorsal orifice, the foramen ovale. Thus, the dorsal and cranial rims of this orifice are formed by the

septum secundum, whereas the ventral and caudal rims are a contribution of the myocardialized atrial spine.

In humans, the septum secundum is difficult to identify during embryonic development, becoming evident during the second trimester (Lamers and Moorman 2002). The atrial septa forms part of the mediastinal component of the atrial tissue, along with that derived from the pulmonary vein.

During development, there is a pressure gradient between right and the left atria (in that direction) that allows the flow of blood between the two chambers through the foramen ovale, displacing the septum primum to the left and passing though the septum secundum. This is the flap valve of the oval foramen.

Normally, during and shortly after birth, when the flow through the umbilical vessels stops and the pulmonary circulation expands, the pressure gradient is inverted and the septum primum is tightly applied against the septum secundum, obliterating the atrial shunt.

1.1.2.10 Development of the ventricles

The ventricles derive from the cranial and intermediate limbs of the looped heart tube. During looping, the cranial limb of the loop, the primordium of the conotruncus, receives migrating cells from the cardiac crest and the anterior and secondary heart fields. As described above, after looping, the

inner and outer curvatures carry out different and sequential roles in development (Figure 1.3).

The apical portions of the two ventricles balloon from the outer curvature of the intermediate limb of the cardiac loop. The inlet portion of the limb gives rise to the apical part of the left ventricle, whereas the right ventricle originates from outlet portion, according to the position of the respective primordia in the heart tube (Lamers and Moorman 2002). The myocardium formed by these ballooning events can be distinguished by their expression of working myocardium markers.

While the ballooning of the ventricles start, they are located in sequence, with the primitive atrium connected directly and exclusively with the left ventricle through the developing atrioventricular canal, the left ventricle in turn empties in the right ventricle through the interventricular foramen, surrounded by the primary ring which can be specifically identified by its expression of markers like the G1N epitope (Lamers et al. 1992) and the right ventricle drains in the outflow tract. The outflow tract is entirely supported by the right ventricle (see Figure 1.6).

The wall of the developing right atrium is initially continuous with the wall of the outflow tract through the inner curvature. In order to allow the aligment of the atrioventricular canal to both developing ventricles, the ventricular segment of the inner curvature must undergo an extensive remodelling, mostly accomplished by the expansion of the atrioventricular canal and the displacement of tissue from its right side to the vestibule of the tricuspid valve (Kim et al. 2001).

Several studies in model organisms have begun to reveal the genetic basis of ventricular development (Lyons 1996). The helix-loop-helix (bHLH) transcription factors dHAND/HAND2 and eHAND/HAND1 are expressed predominantly in the primitive right and left ventricular segments, respectively, during mouse heart development (Srivastava et al. 1995; Srivastava et al. 1997). Deletion of dHAND in mice results in hypoplasia of the right ventricular segment. eHAND has also been implicated in left ventricular development, although early placental defects of eHAND mutant mice precluded a detailed analysis of its role in the heart (Firulli 1998).

Mice lacking Nkx2.5 also show lethal defects in ventricular morphogenesis and fail to express eHAND in the heart, suggesting that eHAND may act downstream of Nkx2.5 to control left ventricular development (Biben 1997b). The ventricular-specific homeobox gene Irx4 is dependent on dHAND and Nkx2.5 for expression, and is sufficient, when missexpressed

in the atria, to activate ventricle-specific gene expression (Bao et al. 1999; Bruneau et al. 2000).

In the zebrafish, which has a single ventricle, only one HAND gene (dHAND) has been identified, mutation of which abolishes the ventricular segment of the heart (Yelon et al. 2000). Mice lacking the transcription factor MEF2C, which is normally expressed throughout the atrial and ventricular chambers, also show hypoplasia of the right and left ventricles, resulting in their early demise in embryogenesis (Lin et al. 1997).

1.1.2.11 Septation of the ventricles

The septation of the ventricles is the last step to separate the aortic and pulmonary outflow from the heart cavities, in addition to separate the bicuspid and tricuspid valves, and thereby, in and outflow. The septum has two main components, a membranous part and a muscular part. The muscular component comes from the ventricular wall itself and it grows towards the fused endocardial cushions to finally meet the membranous part. The latter is formed by the union of the fused endocardial cushions with the fused truncoconal ridges (Anderson et al. 2003a). Because of the complexity of its development, the membranous part is where most of the VSD arise. See figure 1.5.

1.1.2.12 Formation and Septation of the outflow tract

The symmetric division of the outflow tract into the systemic and pulmonary outlets starts during the fifth week of normal embryonic development. The outflow tract can be first identified as a purely muscular structure, which connects the embryonic right ventricle with the aortic sac. Later in development, the endocardial jelly inside the tract concentrates to form pairs of cushions facing each other (Anderson et al. 2003b). The cushions form continuous structures running through the length of the outflow tract (ridges), which spiral round one another and solidify. Afterwards, there is an invasion of mesenchymal cells from the neural crest, which enter in a distal to proximal manner and induce the downwards growth and fusion of the ridges to form the aortico-pulmonary septum (Lamers and Moorman 2002).

Partition of the truncus arteriosus will also contribute to the formation of the aortic and pulmonary semilunar valves.

1.1.2.13 Development of the valves

During development, division of the cardiac tube in cavities is achieved by means of regional swellings of extracellular matrix, called cardiac cushions that mark the location of the valves. Reciprocal signalling between the endocardial and myocardial cell layers in the cushion region, mediated in

part by TGF-* family members, induces a transformation of endocardial cells into mesenchymal cells (Brown et al. 1999). These migrate into the cushions and differentiate into the fibrous tissue of the valves; they are also involved in septation of the common atrioventricular canal into right- and left-sided orifices. Studies using gene targeting in mice has revealed novel roles for the NF-ATc and Smad6 transcription factors in the formation of cardiac valves. Mice lacking NF-ATc, a downstream mediator of signalling by the calcium-dependent protein phosphatase calcineurin, exhibit fatal defects in valve formation, reflecting a potential role for calcineurin in transduction of signals in valve development (de la Pompa et al. 1998; Ranger et al. 1998). Galvin et al. showed that Smad6, which is implicated in the activation of gene expression in response to TGF-f signalling, is also expressed specifically in cardiac valve primordia, but disruption of Smad6 leads to abnormal valves (Galvin et al. 2000).

1.2 The myosins

The myosins are a family of proteins with members present in a wide range of organisms, from yeast to plants and humans. They have in common the ability to produce movement through interaction with actin and hydrolysis of ATP (Mermall et al. 1998). In general, they display a modular structure comprising a motor domain, a neck domain that can bind light chains or calmodulin and a tail domain that allow the protein to form complexes with other myosin molecules, for example the thick

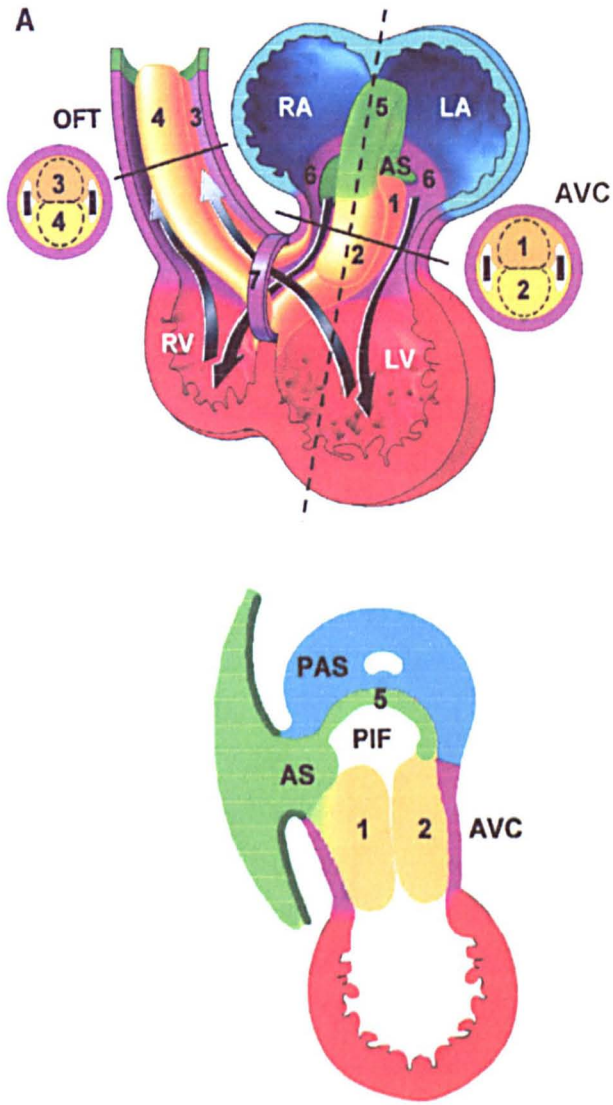


Figure 1.6: A. Diagram showing a coronal (A) and sagital (B) sections (corresponding to dashed line in A) of the heart just before septation. 1) Inferior atrioventricular endocardial cushion; 2) Superior atrioventricular endocardial cushion; 3) Parietal outflow tract ridge; 4) Septal outflow tract ridge; 5) Spur of the atrial spine on the edge of septum primum; 6) Non-balloned portion of atrial heart tube; 7) Primary ring. The atrial appendages (blue) and ventricles (red), are ballooned from the outer curvature and non-ballooned portions of the primitive heart tube (purple) later contribute to septation. AS: atrial spine, PAS: septum primum; PIF: ostium primum; RA: right atrial appendage; LA: left atrial appendage; LV: developing left ventricle; RV: developing right ventricle. The circles show transversal sections of the atrioventricular canal (AVC) and outflow tract (OFT) showing the orientation of the endocardial cushion and ridges. Figure taken from: Moorman AFM, Lamers WH: Cardiac septation: a late contribution of the embryonic primary myocardium to heart morphogenesis. Circ Res. 2002;91:93-103.

filament of the sarcomere, in order to be in position to interact with actin (Sellers 2000).

Myosins are phylogeneticaly classified in at least 18 classes according to their structure. In humans members of 11 classes have been found: I, II, III, V, VI, VII, IX, X, XV, XVI, XVIII (Berg et al. 2001). It has been shown that mutations in these genes can cause human disease (see table 1.1).

Table 1.1: Human diseases caused by mutations in myosin genes

Class	Symbol	Chrom.	Fenotype	Reference
I	*MYO1A*	12	Neurosensory Deafness	(Donaudy et al. 2003)
II	*MYH2*	17	Inclusion Body Myopathy	(Martinsson et al. 2000)
	MYH6	14	Atrial Septal Defect	(Ching et al. 2005)
			Hypertrophic Cardiomyopathy	(Niimura et al. 2002)
	MYH7	14	Hypertrophic Cardiomyopathy	(Anan et al. 1994)
			Laing Distal Myopathy	(Meredith et al. 2004)
	MYH8	17	Carney Complex	(Veugelers et al. 2004)
	MYH9	22	May-Hegglin Anomaly	(Seri et al. 2000)
			Sebastian Syndrome	
			Fetchner Syndrome	
			Neurosensory Deafness	(Lalwani et al. 2000)
	MYH14	19	Neurosensory Deafness	(Donaudy et al. 2004)
III	*MYO3A*	10	Neurosensory Deafness	(Walsh et al. 2002)
V	*MYO5A*	15	Griscelli Syndrome	(Pastural et al. 1997)
VI	*MYO6*	6	Neurosensory Deafness	(Melchionda et al. 2001)
VII	*MYO7A*	11	Usher Syndrome	(Weil et al. 1995)
			Neurosensory Deafness	(Liu et al. 1997)
XV	*MYO15A*	17	Neurosensory Deafness	(Wang et al. 1998)
MLC	*MYL3*	3	Hypertrophic Cardiomyopathy	(Poetter et al. 1996)
	MYL2	9	Hypertrophic Cardiomyopathy	

1.2.1 Class II myosins

Myosins of class II were the first ones to be described. Because of that they are called "conventional myosins". These myosins form hexameric complexes: two heavy chains of molecular weight ranging from 100,000 Da to 250,000 Da and two pairs of light chains between 15,000 Da and 20,000 Da (Korn 2000).

The amino-terminal end or "head" of the class II myosin heavy chain contains the motor domain, adjacent to the "neck" that features two IQ domains. The carboxyl end of the molecule consist of a long coiled-coil tract which homodimerizes with another myosin heavy chain forming the "rod", giving the typical two-headed structure of this complexes which allows them to form filaments (Hodge et al. 1992; Rayment et al. 1993b).

According to the phylogenetic sequence of the motor domain, class II myosins can be classified in four different categories: 1) sarcomeric myosins from striated skeletal and cardiac muscles, 2) vertebrate smooth and non-muscle myosins, 3) myosins from protozoans and other lower eukaryotes and 4) myosins from fungi (Sellers 2000).

The motor domain is constituted of four subdomains: 1) An amino-terminal subdomain, displaying a SH3-like motif.; 2) The upper 50 kDA subdomain which shows in its surface the so called "hypertrophic cardiomyopathy (HCM) loop" because missense mutations changing amino acid residues

located in it can cause the disease (Geisterfer-Lowrance et al. 1990); 3) The lower 50 kDa subdomain and 4) The converter region from which a long a helix emerges to form the neck region which has two IQ domains. The essential light chain binds to the closest to the converter region whereas the regulatory light chain binds to the second.

These submotifs are linked by flexible tracts that allow certain mobility between them. They are two protease sensitive parts of the myosin motor domain whose structure using crystallography as not been resolved, probably because of their flexible nature (Mornet et al. 1981): 1) Loop 1, close to the ATP binding site and 2) Loop 2, in the actin binding site. These two regions are hypervariable amongst class II myosins and interestingly they are also subject of alternative splicing (Itoh and Adelstein 1995).

The upper and lower 50 kDa subdomains form most of the actin binding site. As a three-dimensional structure, these two subdomains are separated by a cleft which closes slightly when binding ATP and closes even more when binding actin molecules (Houdusse et al. 1999).

The ATP binding site is an open pocket in which three conserved sequence elements can be recognized: the Switch 1 (residues 233-237), Switch 2 (Asp 454) and P-loop (residues 181-187) (Minehardt et al. 2002). Their functions are not yet clear. The cavity is formed mainly by seven β-strands (Rayment et al. 1993b).

1.2.2 Cardiac Myosin Heavy Chains

Cardiac myosins are the main cytoskeletal component and force generator molecules of the heart. Myosin complexes in the ventricles are designated V_1, V_2 and V_3. V_1 is formed by two α-cardiac myosin heavy chains (α-MHC), V3 by two β-cardiac myosin heavy chains (β-MHC) and V2 by one of each type (Morkin 2000).

The ventricular content of myosin complexes varies according to the speed of contraction. Fast-contracting ventricles like those of small rodents contain mainly of V_1, those of intermediate speed (for example rabbit or guinea pig) are mostly formed by V_3 and lower amounts of V_1 and V_2 whereas slow ventricles (human and bovine) contain less than 10% of V_1 (Swynghedauw 1986).

The amino acid sequences of mouse α- and β-MHC are 93% (Krenz et al. 2003) similar, whereas the similarity between their human counterparts is 92%.

It has been shown that the rabbit V_1 myosin complex has a 2-3 fold faster actin filament sliding velocity compared with V_3 but generates only about 50% the isometric force (Palmiter et al. 1999). It is thought that these functional differences between α- and β-MHC are due to structural differences in the Loop 1, located between residues 213 and 223 and the

Loop 2 regions, between positions 624 and 646 (Rayment et al. 1993a). Loop 1 appears to modulate velocity through ADP release while Loop 2 regulates the actin dependent ATP hydrolysis rate (Murphy and Spudich 2000).

The genes encoding the human α-MHC and β-MHC are *MYH6* and *MHY7* respectively. They share a remarkable sequence similarity and are located next to each other head to tail in the long arm of chromosome 14 (14q12), being *MYH6* centromeric to *MYH7*. They are separated by 4.5 kb of intergenic sequence and each one contains 39 exons and spans approximately 26 kb of genomic sequence.

The genomic organization of these genes is similar to that found in every placental mammal studied so far (Ensembl Genome Browser, EMBL-WTSI) suggesting that both genes were originated by a duplication in tandem occurred at least before the divergence of eutherians and metatherians, some 125 million years ago (Luo et al. 2003).

1.2.2.1 Control of the Expression of *MYH6*

The expression of cardiac myosin heavy chain begins between 7.5 and 8 days after fertilization at the heart tube stage along with atrial and ventricular myosin light chains and skeletal and cardiac α-actins (Lyons et al. 1990). As the formation of the cavities progresses, by the 10th day of development, the expression of β-MHC is limited to ventricular

cardiomyocytes. At the same time, the levels of α-MHC in these cells decrease while its expression remains high in atrial cells. The α-MHC : β-MHC ratio in ventricular cells is reverted briefly shortly after birth probably due to a surge of the blood levels of thyroid hormone (Lompre et al. 1984). In the human adult β-MHC is the main ventricular myosin whereas α-MHC is predominantly expressed in the atria (Kurabayashi et al. 1988).

The intergenic sequence between *MYH7* and *MYH6* contains a promoter that drives the expression of *MYH6* and an antisense transcript that modulates the expression on *MYH7*. Comparisons of the human and rat *MYH6* promoters show that a portion of sequence between positions –340 and +20 might be sufficient for expression (Molkentin et al. 1996).

Within this segment, several binding targets for transcription factors and modulators have been described. Most of them have been characterized in rodents and later either for experiment or alignment their human counterparts identified (see Figure 1.7). In the human sequence, from downstream to upstream from the transcriptional start, after the TATA box, are found: a proximal CArG site that binds serum response factor (SRF) (Molkentin et al. 1996), two contiguous thyroid hormone response elements (TRE) (Izumo and Mahdavi 1988), a distal CArG site binding SRF (Molkentin et al. 1996), a site that binds an A-rich binding factor (Molkentin and Markham 1994), an M-CAT site for transcriptional enhancer factor 1(TEF-1) (Farrance et al. 1992), two GATA binding sites for GATA-4 (Molkentin et al. 1994), two sites for Ku, a protein from a DNA

binding family of unclear function (Sucharov et al. 2004), a site for myocyte-specific enhancer-binding factor-2 (MEF-2) (Molkentin and Markham 1993), and finally, two sites for Ying Yang 1 (YY1), an ubiquitously expressed transcriptional modulator (Mariner et al. 2005).

Three E-box sites were described in rodents but not found in humans (Molkentin et al. 1996; Navankasattusas et al. 1994). Alignments of the promoter region of the genes encoding α-MHC in different species show several sites of localized conservation besides those mentioned above that are highly likely to be associated in the future with yet undiscovered transcription modulators (see Figure 1.7).

The *MYH6* cis-regulatory sequences are not limited to the conventional promoter. A purine-rich negative regulatory element (PNR) has been characterized in the first intron of the gene (Gupta et al. 1998). It appears to contain two Ets sites that bind PURα and PURβ proteins (see Figure 1.7) (Gupta et al. 2003).

Due to the observation that in normal conditions as well as in pathologic states the level of α-MHC mRNA is not directly related to the levels of protein (i. e. the α-MHC content represents approximately 7.2% of total MHC protein and 35% of the MHC mRNA in healthy subjects), it is believed that the expression of this gene is also subject to translational regulation (Miyata et al. 2000). This regulation seems to be carried out at least in part by creatine kinase BB isoform (Vracar-Grabar and Russell

2004), an enzyme that binds selectively to the 3'UTR of the α-MHC mRNA, localizes in the Z-band of the sarcomere, and is expressed in the embryonic heart and in the adult cardiac tissue in response to ischaemia, hypertrophy and heart fairlure, (Ingwall 2002; Ritchie 1996) known causes of α-MHC downregulation.

Indeed, as well as being regulated by transcription modulators, physiological, pathological and environmental stimuli can modify the expression of *MYH6.* Increased plasma levels of thyroid hormone can up regulate the mRNA levels of α-MHC in the ventricles at any stage of adult life whereas hypothyroidism (where β-MHC can be up regulated) has the opposite effect (Morkin et al. 1983) (Figure 1.7).

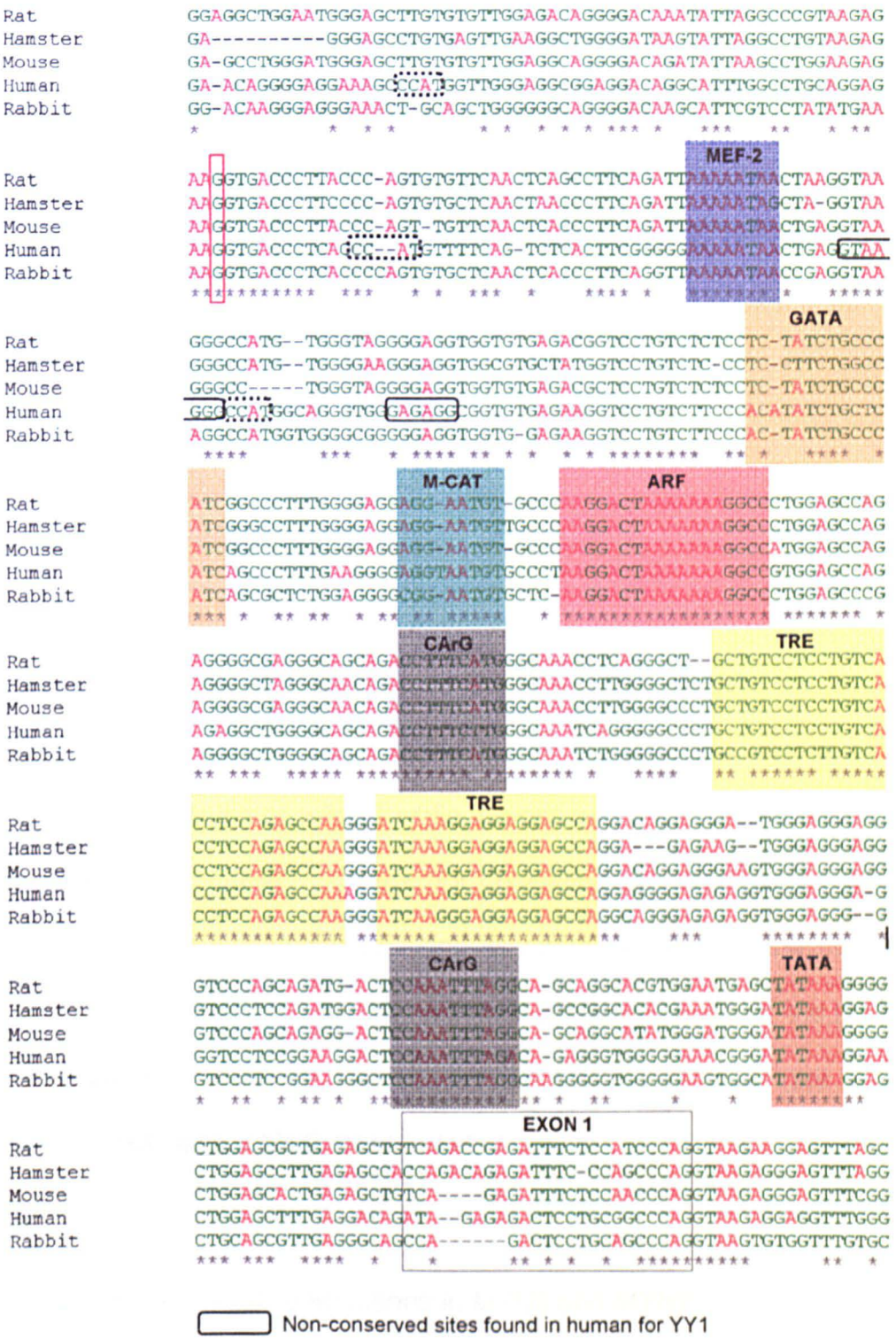


Figure 1.7: Conserved cis-regulatory elements upstream of the transcriptional start of *MYH6*.

The increment in the tension of the heart cavities walls due to pressure or volume overload can also produce up regulation of β-MHC and down regulation of α-MHC (Imamura et al. 1991). Nevertheless, when these conditions are recreated in rodent models, once the stimulus stops, the levels of α-MHC protein are recovered rapidly while the high levels of β-MHC persist for weeks, suggesting that the overload triggered signals regulate both genes independently (Gupta and Gupta 1997).

These differential and in some instances antithetical expression profiles of the *MYH6* and *MYH7* genes are thought to be controlled mainly by the regulatory sequence located between both genes which works as a bi-directional promoter, formed by two clusters of transcription modulators binding sites. The one which is closer to the transcriptional start of *MHY6* drives the expression of α-MHC while the one nearest to *MYH7* drives the expression of an antisense transcript originating 2kb downstream of the *MYH7* gene and is transcribed along its entire length from the opposite strand. This transcript is believed to prevent the processing of the β-MHC pre-mRNA and could be important to regulate rapid changes in the expression of cardiac MHC (Haddad et al. 2003).

1.2.2.2 Disease Causing Mutations in *MYH6* and *MYH7*

Because of their sequence similarity and close vicinity, *MYH6* and *MYH7* are potentially prone to non-homologous recombination events. The outcome of such an event has been reported in a family with hypertrophic

cardiomyopathy (Tanigawa et al. 1990). The resultant hybrid *MYH6/MYH7* gene has a 5' portion identical to *MYH6* that ends around exon 27 where it continues with a 3' segment identical to *MYH7* until the end of the gene. This hybrid is located between two complete copies of *MYH6* and *MYH7* and it was thought to be the cause of the phenotype in that family. A missense mutation in the normal copy of MYH7 on the same chromosome (Arg4532Cys) was discovered later in that family as well as in other hypertrophic cardiomyopathy families without the hybrid gene (Watkins et al. 1992), therefore concluding that the rearrangement probably does not have a functional significance.

Since then, it has been estimated that nearly half of hypertrophic cardiomyopathy cases are caused by missense mutations in *MYH7* (Watkins et al. 1992) and mutations of the same gene producing dilated cardiomyopathy (Kamisago et al. 2000), Laing distal myopathy (Meredith et al. 2004) and myosin storage myopathy (Tajsharghi et al. 2003) have been also described.

In humans, only two disease causing mutations of *MYH6* have been found. Arg795Gln has been related to hypertrophic cardiomyopathy of the elderly (age of onset 74). This mutation replaces a hydrophilic glutamine for a basic arginine in a position located in the first IQ domain, possibly disrupting the interaction with the essential myosin light chain (Niimura et al. 2002). Ile820Asn was found in a large family with autosomal dominant atrial septal defect and is located in the second IQ domain, where has

been shown to prevent the interaction with the regulatory light chain (Ching et al. 2005).

1.2.2.3 Animal Models of Mutations in *MYH6*

Homozygous transgenic mice for null *Myh6* alleles die in utero at 11-12.5 days post-fertilization displaying gross heart defects while the heterozygotes, though appearing externally normal, show a decrease in the length and disruption of the cardiac sarcomere pattern with wide Z-bands and shortening of the A-band. The ventricles show significant reductions in contractility and relaxation parameters and multifocal fibrosis. These changes however had incomplete penetrance (5/8) probably due to genetic or epigenetic variability or to still unknown compensatory mechanisms (Jones et al. 1996).

Another murine model of *MYH6* mutation was constructed with a modification identical to one of the mutations observed in human *MYH7* (Arg403Gln) in cases of hypertrophic cardiomyopathy. The homozygous α-$MHC^{Arg403Gln/Arg403Gln}$ mice die by a fulminant dilated cardiomyopathy at the 8^{th} of postnatal life due to myocyte loss, showing ventricular multifocal transmural necrosis and secondary calcification. The sarcomere structure was normal apart from some focal myofibrillar disarray. Heterozygous α-$MHC^{Arg403Gln/+}$ mice develop myocardial histologic abnormalities similar to those found in cases of familial hypertrophic cardiomyopathy (Fatkin et al. 1999).

During a screening for Zebrafish mutations affecting cardiac chamber formation, Berdougo et al. (Berdougo et al. 2003) described the locus *weak atrium* which later was found to encode the Zebrafish atrial myosin heavy chain. *weak atrium* mutants show atrial contractility defects, the atrial chamber becomes dilated, a characteristic blood pool caudal to the atria develops and the atrial sarcomere shows disorganization of myofilaments. These mutants do not exhibit abnormalities outside the heart and can become fertile adults probably because the *weak atrium* protein product is not expressed in the ventricles.

1.3 Congenital heart defects

1.3.1 Aetiology

The aetiology of congenital heart defects remains unknown in most cases. Approximately a quarter of cases of CHD occur as part of a complex with malformations in other organs in the form of an association, Mendelian syndrome or a chromosomal abnormality (Ferencz 1993; Ferencz et al. 1989). The remaining three quarters of cases occur as isolated malformations, mainly sporadically. Both environmental and genetic factors seem likely to be implicated in these cases (Nora 1993).

1.3.2 Identified environmental factors

In newborns from mothers infected by rubella during pregnancy the risk of CHD is 35% (Burn 2002), the more common being persistent ductus arteriosus, pulmonary arterial stenosis and septal defects (Emmanouilides et al. 1964; Rowe 1963).

Maternal diabetes has long been known to increase the incidence of CHD to 4%, approximately five times the incidence in the general population. Ventricular septal defect, coarctation of aorta and complete transposition were found in more than 50% of the cases (Rowland et al. 1973). Different degrees of the left isomerism sequence have been also reported (Splitt et al. 1999).

Early diagnosis and dietary restrictions during childhood have allowed female patients with phenylketonuria to reach reproductive age. Unless a previously abandoned phenyalanine-free diet is reintroduced before conception, the risk of CHD for the unborn child increases to between 25% and 50% (Burn 2002; Walter 1995), especially tetralogy of Fallot (Lenke and Levy 1980).

It has been estimated that fetal alcohol syndrome affects 1% to 3% of newborns (Mengel et al. 2006), approximately 30% to 40% of them additionally show a CHD, typically ventricular or atrial septal defect and tetralogy of Fallot (Moss 1992).

Maternal consumption of the anticonvulsant drugs like valproates and hydantoin has been related to CHD. Fetal hydantoin syndrome occurs in 10% of exposed pregnancies (Holmes et al. 2001) and a fraction of them suffers a form of CHD, but no specific lesion seem to predominate (D'Souza et al. 1991; Lin 1990).

Lithium exposure during gestation has been linked to Ebstein's malformation (Weinstein and Goldfield 1975), whereas according to other reports the incidence of a particular CHD does not prevail (Kallen 1987).

The incidence of CHD in monozygotic twins is considerably higher than in dizygotic twins or products of single pregnancies (Burn and Corney 1984), reaching 2.3% in some estimates (Karatza et al. 2002), without considering those cases where twin-twin transfusion syndrome occurs. For monochorionic-diamniotic monozygotic twins the risk of a CHD has been calculated as 7%, and 57% for monochorionic-monoamniotic twins

(Manning and Archer 2006), ventricular septal defects and right isomerism being the more common anomalies.

Products of pregnancies conceived by *in vitro* fertilization show a fourfold incidence increase of CHD in both singletons and twins (Koivurova et al. 2002).

Usually accounting for 5% to 10% of CHD, patent ductus arteriousus is the more common CHD in children born at high altitudes like Mexico City or the Peruvian Andes where the incidence of this defect is 30 times the observed at sea level (Fyler 1980; Penaloza et al. 1964), possibly due to decreased levels of blood oxygen. The same mechanism is likely to influence the high incidence of the same defect in premature infants (Ellison et al. 1983).

1.3.3 Atrial Septal Defects

Atrial septal defects are amongst the most common congenital cardiac defects. The mildest forms of the disease can be sub-clinical during many years. However, blood circulation consequences develop along time. The higher blood pressure from the left atrium causes blood to pass to the right

atrium, which leads to pulmonary hypertension. Over the years, this increases the pressure on the right ventricle, the right atrium and ultimately causes reversal of flow in the interatrial communication, mixing both oxygenated and non-oxygenated blood, which is translated into cyanosis in the patient. On the other end of the spectrum of septal defects is the premature closure of the foramen ovale during development. When this occurs prenatally, the blood flow passes completely from the right atrium to the right ventricle, causing considerable hypertrophy of the right side of the heart. In contrast, the left side is hypoplastic due to the substantial decrease of blood flow. Individuals suffering this condition typically die during the first days of life.

A persisting communication between the two atria (shunt) can be caused by a range of events during early development. The most common varieties of this malformation are due to hypoplastic growth of the septum secundum or lack of resorption of tissue around the foramen secundum. On the other hand, a lack of union between the leading edge of the septum primum and the endocardial cushions can give raise to a lower less common defect. If the malformation involves the endocardial cushions, the atrioventricular valves can be affected, making the clinical outcome more severe. The presence of a common atrium is one of the most serious conditions of this kind, and it is produced by lack of septation of the atria.

1.3.4 Tricuspid atresia

This defect consists in the complete occlusion of the opening between the right and left ventricle. The direct consequence of this would be death by lack of blood oxygenation, as the blood can not get access to the lungs. However, the patients can survive during months or years thanks to the existence of secondary defects or shunts. In order to allow the blood to reach the lungs from the right atrium, an interatrial defect must exist so the blood can pass to the left atrium and then to the left ventricle. Secondly, one or more secondary shunts must permit blood to circulate to the pulmonary arterial system, either by an interventricular septal defect or from the aorta through a patent ductus arteriosus. From the lungs, the blood goes back to the left atrium and could be recycled through the lungs before reaching the systemic circulation.

1.3.5 Ventricular septal defect

VSD is the most common septal defect. The membranous part of the interventricular septum is a weak spot due to the convergence of several embryonic tissues at this level. Consequently, in 70% of the cases the VSD lies in this area. In the initial stages of the disease, the higher left ventricular pressure leads to a left to right shunt without cyanosis.

However, right ventricular hypertrophy can develop, causing pulmonary hypertension and a reversal of the shunt. Notably, in the majority of cases the defect closes spontaneously before age of ten.

1.3.6 Persistent truncus arteriosus

During normal embryonic development, the outflow tract is divided in order to generate the pulmonary trunk and the aorta. However, in some individuals the truncoconal ridges do not fuse to form the septae that should divide the tract, and thus there is a persistence of a large single outflow tract overriding the ventricular septum. This outflow vessel receives the blood flow from both ventricles and it is almost always accompanied by a ventricular septal defect. This condition is clinically translated into severe cyanosis and death during the first year of life, if not corrected.

1.3.7 Transposition of the Great Vessels

This condition is the most common cause of cyanosis in the newborns. When the truncoconal ridges fail to spiral when dividing the outflow tract, two independent circulatory systems are created, with the right ventricle

emptying venous blood into the aorta and the general circulation, and the left ventricle pumping highly oxygenated blood to the lungs through the pulmonary artery. This malformation is compatible with life only when associated with ASD, VSD and patent ductus arteriosus.

1.3.8 Tetralogy of Fallot, Aortic and Pulmonary Stenosis

Aortic and pulmonary stenosis may be caused by asymmetric partitioning of the outflow tract. Complete occlusion of one of the vessels is called aortic or pulmonary atresia.

One example of asymmetric fusion of truncoconal ridges is the Tetralogy of Fallot (TOF), which is the most common cause of cyanosis in infants. TOF consists in pulmonary stenosis, a large aorta, right ventricular hypertrophy and membranous interventricular septal defect. Poorly oxygenated blood is pumped through the enlarged aorta to the circulatory system and causes cyanosis from birth. Clubbing of the fingers is also a common feature due to chronic hypoxemia. The prognosis is poor, with a survival rate of 50% if not corrected within the first two years of life.

1.3.9 Patent Ductus Arteriosus

During normal postnatal development, the ductus arteriosus is spontaneously closed, separating the aortic arch from the pulmonary artery. If this structure persists, a slow progressive disease occurs. The condition is silent at birth, but over the years, the high aortic pressure causes blood to flow towards the pulmonary circulation through the patent ductus leading to pulmonary hypertension, and eventually, heart failure. Lower-body cyanosis can appear due to an inversion of the blood circulation through the ductus. There is a higher incidence of patent ductus arteriosus in pregnancies complicated by rubella infection or hypoxia.

1.3.10 Coarctation of the Aorta

The embryogenesis of the coarctation of the aorta is not well known yet. Several factors could lead to the same malformation, for example, there is a higher incidence of coarctation of the aorta in Down and Turner's syndrome patients. This anomaly consists of a narrowing of the descending aorta, which could be upstream or downstream from the ductus arteriosus. Usually, the ductus arteriosus remains open in the preductal type, and the blood reaches the descending aorta through it.

However, there is also flow of venous blood through the ductus, which causes cyanosis of the lower trunk and limbs.

The most common form (95% of cases) is the postductual coarctation. As the arterial circulation of the upper and lower part of the body is separated by the narrowing, collateral circulatory channels between small arteries are developed to communicate both systems. As a consequence, the small arteries carry an abnormally large blood flow, which clinically translates as notches in the ribs by radiographic analysis and higher blood pressure in the arms than in the legs.

1.4 Recurrence risk in siblings and offspring of CHD patients.

A previous antecedent of a CHD case in the family was the single largest determinant in the CHD cases cohort of the large Baltimore-Washington study (Boughman 1993; Boughman et al. 1987). Even though this points to an important genetic component in the aetiology of CHD, several studies have reported a very different empirical offspring recurrence risk of the disease in general and specific lesions (Czeizel et al. 1982; Emanuel et al. 1983; Whittemore et al. 1982). This disparity was probably due to selection bias. When the bias is corrected the recurrence risk for offspring is 4.1% which is significantly higher than the recurrence risk for siblings

(2.1%; p=0.021) (Burn et al. 1998). During the same study, when the data for individual defects were used in computer inheritance pattern modelling, atrioventricular septal defect appeared to be a single gene disorder, tetralogy of Fallot as an oligogenic defect and transposition as an entirely sporadic entity. The ratio of affected offspring from affected mothers versus those from affected fathers was 2.23. This phenomenon could be due to selective imprinting of the allele from the father and dependence of the maternal allele for normal function.

Chapter 2 MATERIALS AND METHODS

2.1 Bacteriological techniques

2.1.1 Preparation of calcium chloride competent cells

An isolated, single colony was picked using a sterile toothpick from a plate previously incubated for 16-20 hours at 37°C. Material from the colony was transferred into 5 ml of LB medium and incubated in rotation overnight. Two ml were taken from the overnight culture and used to inoculate 100 ml of plain LB medium which in turn is incubated in rotation from 90-120 minutes, until optical density reaches OD_{600}=0.6.

When the desired OD was achieved, the cultured was harvested by centrifugation at 3500rpm for 15 minutes at 4°C, the resultant pellet re-suspended in 50ml of ice cold 50mM $CaCl_2$ and 10% glycerol, snap frozen with liquid nitrogen and stored in 100 µl aliquots at -80°C.

2.1.2 Preparation of electrocompetent cells

Overnight cultures in 400ml of LB were centrifuged at 4000g for 10

minutes. The resultant pellet was re-suspended in 200ml of ice cold sterile distilled water. The suspension was centrifuged again at 4000g for 10 minutes and the new pellet re-suspended with 100ml of ice cold sterile distilled water and the process repeated with 10ml of ice cold 10% glycerol and finally with 1ml of 10% glycerol. The suspension was distributed in 50µl aliquots and these frozen in liquid nitrogen and stored at -80°C.

2.1.3 Transformation of calcium chloride competent cells

Aliquots of $CaCl_2$ competent cells were thawed on ice and a solution of DNA at the appropriate concentration was applied and mixed using a chilled pipette tip. The mix was incubated for 30 minutes on ice and then transferred to a 42°C water bath and incubated there for 90 seconds. The tube was transferred again to ice, to be incubated for 2 minutes and then applied to a loosely capped tube with 5ml of SOC medium where the reaction was incubated for 60 minutes in a shaker incubator at 150 rpm. The culture was harvested in a tabletop centrifuge at 8000rpm and the pellet re-suspended in 200µl of LB medium. 100 µl of the suspension was applied to LB-agar plates with the appropriate selective reagent. These plates were incubated overnight at 37°C.

2.1.4 Transformation of electrocompetent cells

A 50µl aliquot of electrocompetent cells was thawed on ice and a solution

of DNA applied and mixed. The mixture was then transferred to a pre-chilled electroporation cuvette (Eurogentec) and incubated in ice for 30 minutes more. Next, the cuvette was dried externally and located in a GenePulser (Biorad), where the electroporation takes place. The following GenePulser parameters were used: capacitance, 25μFD; capacitance extender, 125μFD; resistance 200Ω; and voltage 2500V. After electroporation, the mix was applied to a tube with 750μl of SOC medium and put in shaker incubator for 60 minutes at 37°C. The suspension was then spread in LB-agar Petri dishes at variable volumes.

2.1.5 Preparation of selective media

LB-broth and LB-agar were supplemented with antibiotics as a selection measure. In the case or the agar the substance was added to the molten medium just before was poured and allowed to set in Petri dishes. The final concentration of each agent on the media was: ampicillin 100mg/l, kanamycin 25mg/l.

2.1.6 Preparation of glycerol stocks

Glycerol stocks were prepared adding 300μl of 80% glycerol solution to 500μl of overnight broth culture. The mix was snap frozen in liquid nitrogen and stored at -80°C.

2.2 Nucleic acid amplification (Polymerase chain reaction)

2.2.1 RNA extraction

Animal tissue was obtained from 11.5 dpc mouse embryos, dissecting limbs and hearts and flash freezing them in liquid nitrogen. RNA was extracted following the protocol for Purescript Total RNA Purification Kit (Gentra). All the glass and plastic ware used were treated with DEPC (diethyl pyrocarbonate) water to avoid RNase activity. In brief, 5 to 10 mg of fresh animal tissue were collected in Eppendorf tubes containing 300µl of lysis solution, and homogenized with an RNase-free plastic pestle. 100 µl DNA-protein precipitation solution were added to the lysate and mixed by inverting the tubes. After 5 minutes on ice, the solution was centrifuged for 3 min to separate the RNA from DNA and proteins, which precipitate in a pellet. The supernatant was transferred to a fresh tube and 300 µl of 100% Isopropanol were added to precipitate the RNA. After inverting several times, the RNA was pelleted by centrifugation for 3 min at full speed, washed in 70% ethanol, dried and resuspended in sterile distilled water treated with DEPC.

2.2.2 Reverse transcription

In order to produce cDNA for downstream experiments, total RNA was

reversed transcribed using the 1st strand Synthesis Kit from ABgene. In brief, after heat denaturation at 72°C, 1 µg of RNA is mixed with oligo-dT, reverse transcriptase (blend of AMV and MMuLV), 5X buffer, 5nm dNTPs and RNase inhibitor. The solution was incubated at 45-50°C for 90 minutes, followed by 10 minutes at 75°C for enzyme inactivation. RT reactions were set up without adding reverse transcriptase, to use as negative controls.

2.2.3 Primer design

Design of oligonucleotides to be use as PCR primers was assisted by the web interfaces of the Primer3 program at the Whitehead Institute for Biomedical Research (Rozen 2000) and NetPrimer (Premier Biosoft International).

2.2.4 PCR reactions

PCR reactions were performed in volumes of 20 and 50 µl, according to the amount of product required in the downstream application. All the reactions were carried out in independent 200µl tubes or 96 well plates and several different thermal cyclers. The typical recipe is showed below:

	20 µl reaction	50 µl reaction
Forward primer 10pmol/µl	1 µl	2.5 µl

Reverse primer 10pmol/µl	1 µl	2.5 µl
dNTPs 2mM	2 µl	5.0 µl
10x PCR reaction buffer *	2 µl	5.0 µl
Taq polymerase 5U/µl	0.1 µl	0.5 µl
Sterile distilled water	11.9 µl	30.5 µl
DNA [50ng/µl]	2 µl	5 µl

*100 mM Tris-HCl, pH 8.3 at 25°C; 500 mM KCl; 15 mM $MgCl_2$; 0.01% gelatin

2.3 DNA manipulation techniques

2.3.1 Agarose gel electrophoresis

Standard 1% (w/v) agarose gel was made by melting 0.5g of agarose in a total volume of 50ml 1X TAE (40mM Tris-Acetate, 10mM EDTA) and adding 0.5µg/ml Ethidium bromide. Gel loading buffer (30% glycerol, 0.025% (w/v) Bromophenol Blue, 0.025% (w/v) Xylene Cyanol) was added to DNA samples before loading onto agarose gels.

2.3.2 Gel extraction

DNA was extracted from agarose gels using the Gel extraction kit supplied by QIAGEN according to the manufacturer's instructions. Briefly, the

bands containing DNA were carefully excised from the gel with a sterile scalpel, mixed with 3 volumes of a buffer containing guanidine thiocyanate and heated at 50°C until the gel slice was dissolved. The mixture was applied to a QIAquick silica column and centrifuged. Salts were washed away by applying a buffer containing ethanol to the column and centrifuging. Residual ethanol was eliminated by an additional centrifugation step. The DNA was eluted from the column applying sterile distilled water, letting stand for 5 min and centrifuging at maximum speed.

2.3.3 Digestion of DNA

Basic digests were carried out using the corresponding restriction enzymes, according to the restriction map of the DNA template. All the reactions were performed according manufacturer's instructions.

2.3.4 Ligation

Direct ligation of PCR products were performed using the pGEM-T easy system kit (Promega), according to the manufacturer's instructions. In brief, the reactions are set up adding T4 DNA ligase (3U/µl), T-easy vector (50ng), 2X ligation buffer, the PCR product, and sterile distilled water to the desired volume. The ligation mixture is then incubated overnight at 4°C.

Ligation of DNA fragments produced by restriction digest was performed using T4 DNA ligase (Invitrogen). Reactions were set up adding 5X ligase buffer, insert:vector at a 3:1 molar ratio, T4 DNA ligase 0.1 to 1 units and sterile distilled water to the desired volume. Incubation was carried out overnight at 14°C.

2.3.5 Plasmid DNA isolation

Bacterial cultures were grown overnight and harvested by centrifugation at 3000Xg for 1 minute. Plasmid DNA was isolated by using a QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions.

2.3.6 Measurement of DNA concentration

A 1/10 dilution of the original miniprep was measured in a GeneQuant RNA/DNA calculator (Pharmacia Biotech), using UV wavelength of 260nm, following the manufacturer's instructions.

2.3.7 Nucleic acid precipitation

For plasmid or genomic DNA, 0.7 volumes of isopropanol, or 3 volumes of 100% ethanol with 0.1 volumes of 3M sodium acetate pH=5.2 were added to the DNA solution and mixed thoroughly. For the isopropanol precipitation the mixture was centrifuged at 12000Xg for 15 minutes at

room temperature, whereas for the ethanol precipitation the DNA/ethanol solution was centrifuged for 10 minutes at 12000Xg to recover the DNA. After carefully removing the supernatant, the DNA pellet was washed with 70% ethanol and centrifuged for 2 minutes at 12000g. The supernatant was removed and the excess liquid was evaporated at room temperature. The DNA pellet was dissolved in the desired volume of TE buffer or sterile distilled water.

For primer purification, 15 µl 3M Na acetate (pH 5.6) and 1.5 µl 1M $MgCl_2$ were added to 150 µl 100µM primer and mixed. Afterwards, 500 µl of cold 100% ethanol were added, the mixture was chilled on ice for 15 to 20 minutes and centrifuged for 15 to 20 min. This was followed by an 80% ethanol wash and 5 minutes spin. After removing the supernatant, the pellet was air dried and resuspended in 120µl of sterile distilled water, for a final concentration of 100µM.

2.3.8 End-labelling of DNA oligonucleotides

^{32}P γATP was used to end-label oligonucleotides in the following manner. Approximately 50 pmol of each oligonucleotide were diltuted in 2µl of sterile distilled water. The mixture was heat denaturated by incubating it at 65° for 3 minutes to remove secondary structure in the oligonucleotide. The tube was then placed immediately on ice. Afterwards, 1µl of 10x phosphonucleotide kinase buffer (GIBCO), 1µl of ^{32}P γATP (AMERSHAM) and 1µl of T4 phosphonucleotide kinase (GIBCO) were added, and the

mixture was incubated for one hour at 37°C. The radioactively-labelled PCR products were visualized using a PhosphorImager (Molecular Dynamics).

2.4 Microsatellite typing

2.4.1 Sample preparation

The loading buffer for microsatellite gels was prepared mixing 500µl of 0.5M EDTA and 9.5ml of formamide (SIGMA). Approximately 5 mg of bromophenol blue and xylene cyanol powder were added to the solution to colour it.

3µl of each radioactively labeled PCR product were mixed with 7µl of loading buffer, and 4µl of this solution were loaded on each well on the 8% denaturing polyacrylamide gels, prepared as below.

2.4.2 Radioactive microsatellite gels

The mobility of PCR products was assessed in 8% denaturing polyacrylamide gels. The gels were prepared by mixing 80ml of 40% acrylamide solution with 168g of urea and 80 ml of 5X TBE and incubated at 65°C until completely dissolved. Enough sterile distilled water was added to obtain to 400ml of an 8% solution. To polymerise the gel, 640µl of 10% APS and 80µl TEMED were added to 80ml of this solution.

The gels were run at 55W for 3 hours at room temperature, using a Life Technologies S2 gel electrophoresis system and a BioRad Power Pac 3000. On completion, the gels were dried on a BioRad gel dryer with vacuum pump and examined using a Molecular Dynamics Phosphorimager.

2.5 Northern blot

2.5.1 Radiolabeling of probe

Approximately 25 ng of template DNA were mixed with random hexamer primers, heated for 2 min and placed on ice for 1 min. The solution was spun down at 4°C and returned to ice. 50µl reactions were set up adding 1µl 5 mM dNTPs, 10 µl 5X buffer, 10 mCi/ml [α-^{32}P]dCTP, 5 units of Klenow fragment of E. coli DNA polymerase I and sterile distilled water. The mixture was incubated for 1 hour at room temperature, and the radiolabeled probe was used immediately.

2.5.2 Hybridization

RNA was run on a denaturating agarose gel (1.25g agarose, 87ml sterile distilled water, 10mL MOPS/EDTA and 5.1mL formaldehyde). The RNA was transferred to a nylon membrane and hybridized, essentially as described in the Molecular Cloning Manual (Sambrook 2001).

2.6 Informatic analysis

2.6.1 BLAST

The Basic Local Alignment Search Tool (BLAST) is one of the algorithms available from the NCBI (National Centre for Biotechnology Information) website. The BLAST program allows us to compare nucleotide and peptide query sequences against the public databases, in a local manner. The resulting information is valuable for identification of homologue genes or proteins, evolutionary conservation and classification into families.

2.6.2 GLUE

The Genetic Linkage User Environment (GLUE) is a free software based at the LITBIO (Laboratory for Interdisciplinary Technologies in Bioinformatics) website (formerly at HGMP, Human Genome Mapping Project Resource Centre). GLUE is a graphic interface to programs used for linkage and statistical genetics. Packages currently supported by GLUE are Linkage, Genehunter, Merlin, Unphased, and Transmit. The linkage format pedigree files generated in this study were uploaded to the HGMP file space to be processed.

Chapter 3 MUTATIONAL ANALYSIS OF *MYH6* IN SPORADIC CASES OF CONGENITAL HEART DISEASE

3.1 Introduction

Twin studies have provided a way to estimate the genetic component in the aetiology of complex diseases by comparing the concordance between monozygotic (that share both their entire genome and prenatal exposures) and dizygotic twins (sharing half their genes and gestational environment). A high concordance of a complex disease in monozygotic twins is generally accepted as an argument in favour of a strong genetic component.

As the twinning process itself can cause CHD, the low concordance of the disease in monozygotic twins can lead to underestimation of the genetic contribution in its aetiology.

Mendelian forms of common complex disease may be explained as follows: a) the Mendelian form represents another disease with no causal

relation with the sporadic form but they are clinically indistinguishable; b) most of the individuals in the Mendelian family carry or are exposed to most of the susceptibility conditions so in the affected individuals the Mendelian segregation of the remaining susceptibility allele causes the disease; c) the familial aggregation of cases occurred by chance and does not represent a Mendelian form of the disease.

In some cases, a gene responsible for a Mendelian form of a complex disease also behaves as a susceptibility locus for the sporadic form, as in Hirschsprung disease, caused by mutations of *RET* (Carrasquillo et al. 2002). However there are examples of the opposite situation, where the gene mutated in familial cases have no observable contribution for the more common, sporadic form of the disease as in the case of mutations of *BRCA1* and breast cancer (Pharoah et al. 2002).

Mutations of several genes are known to produce Mendelian form of non-syndromic CHD in humans (Table 3.1). Germ-line mutations on several genes have been shown to cause either syndromic or Mendelian isolated congenital heart disease in humans. A single germ-lime mutation of *MYH6* (I820N) has been shown to be responsible for a familial form of secundum atrial-septal defect (Ching et al. 2005).

Table 1 Mutations in Mendelian CHD

GENE	PHENOTYPE	REFERENCE
BMPR2	Pulmonary hypertension,CHD	(Roberts et al. 2004)
CFC1	double-outlet right ventricle, transposition great arteries	(Bamford et al. 2000)
DTNA	Non-compaction with CHD	(Ichida et al. 2001)
ELN	Supravalvular aortic stenosis	(Curran et al. 1993)
GATA4	ASD	(Garg et al. 2003)
GJA1	Hypoplasic left heart syndrome	(Dasgupta et al. 2001)
JAG1	Tetralogy of Fallot	(Eldadah et al. 2001)
NKX2.5	ASD plus conduction defects, Tetralogy of Fallot	(Schott et al. 1998)
TAZ	Non-compaction	(Bleyl et al. 1997)
ZFPM2	Tetralogy of Fallot	(Pizzuti et al. 2003)
MYH6	ASD	(Ching et al. 2005)
THRP2	Transposition great arteries	(Muncke et al. 2003)
ZIC3	CAVC, DILV, DORV, D-TGA, HLHS, IVC, L-TGA, PA, PAPVR, PS, SVC, TAPVR	(Ware et al. 2004)
CRELD1	AVSD	(Robinson et al. 2003)
NKX2.6	Common arterial trunk	(Heathcote et al. 2005)
LDB3	Left ventricular non-compaction	(Vatta et al. 2003)
*TBX5**	ASD, AVSD	(Reamon-Buettner and Borlak 2004)
*HEY2**	AVSD	(Reamon-Buettner and Borlak 2006)
CITED2	ASD, VSD, Tetralogy of Fallot	(Sperling et al. 2005)
ACVR2B	ventricular inversion with ventricular septal defect, inversion and transposition of the great vessels, pulmonary stenosis, total anomalous pulmonary venous return	(Kosaki et al. 1999)
TBX1	TA IAA	(Gong et al. 2001)

CAVC = complete atrioventricular canal; DILV = double inlet left ventricle; DORV = double outlet right ventricle; D-TGA = D-transposition of the great arteries; HLHS = hypoplastic left heart syndrome; IVC = inferior vena cava; L-TGA = L-transposition of the great arteries; PA = pulmonic atresia; PAPVR = partial anomalous pulmonary venous return; PS = pulmonic stenosis; SVC = superior vena cava; TAPVR = total anomalous pulmonary venous return.
*Somatic mosaicism

3.2 Mutational Analysis using denaturing high performance liquid chromatography

Liquid chromatography is an analytical method used for the separation and quantification of the components of a heterogeneous solution through binding to a solid phase (matrix) and selective elution. Denaturing high performance liquid chromatography (dHPLC) is a variant of this procedure that, by means of an ion pairing (tri-ethyl ammonium acetate, TEAA) and a hydrophobic eluant reagent (acetonitrile ACN), allows the separation of DNA or RNA molecules that differ in size or base composition when passed through a column in solution.

The chromatography columns used in dHPLC are made of a solid matrix formed by C18 alkylated polystyrene-divinylbenzene 3µm beads, creating a strongly non-polar phase. A continuous flow of liquid pass through the column during the whole process. At the time of the injection of the DNA solution (generally a non-purified PCR product), the flow through consists of a mix of 50% "Buffer A" (TEAA 0.1M) and 50% "Buffer B" (TEAA 0.1M, 25% ACN).

The positive ammonium ion of the TEAA molecule interacts with the phosphate groups of the hydrophilic DNA molecule, whereas its triethyl

groups coat the nucleic acid with a hydrophobic outer layer that allows interaction with the solid matrix dependent on the length and charge density of the DNA fragment.

The temperatures to be used for dHPLC mutational analysis must be carefully determined. This is achieved by testing several temperatures to select those that, according to simulation by a subprogram of Navigator (Transgenomic), allow all segments of the amplicon to adopt in one of them a helical fraction percentage between 40% and 99%, the optimum for mutational analysis.

In order to elute the DNA molecules from the column, the ratio of A/B Buffers injected decreases, effectively increasing the concentration of ACN while the concentration of TEAA remains constant (gradient stage).

If a heterozygous variation exists within the amplicon, DNA heteroduplexes will form. As more ACN molecules become available in the solution the hydrophobic interactions between the DNA molecules and the matrix become weaker, heteroduplexes are released from the column faster than homoduplexes.

3.3 Aims of study

The aims of this study were: a) To discover *de novo* mutations of *MYH6* that could explain some of the sporadic CHD cases; b) to identify new Mendelian CHD families, not catalogued a such because of low penetrance, where the segregation of *MYH6* mutations could show a Mendelian pattern; c) to find new *MYH6* genomic variants that could be used in future susceptibility studies.

3.4 Materials and Methods

The DNA samples correspond to 144 Australian patent foramen ovale patients and 380 British patients with various kinds of non-syndromic congental heart disease.

The genomic sequence of the *MYH6* gene was obtained from GeneBank at NCBI website (Accesion number: Z20656) and Ensembl Genome Browser (Gene ID: ENSG00000197616). These were aligned to the sequences of the *MYH6* transcript (GeneBank accesión: NM_002471, Ensembl ID: ENST00000356287) using the web interface of the "BLAST 2 Sequences" program at NCBI in order to define intron-exon boundaries.

To cover the entire coding sequence (39 exons) of the gene, 35 PCR amplicons were designed. In general, each amplicon spanned the complete length of an individual exon plus short segments of flanking intronic sequence to either side, to allow the detection of mutations of splicing regulatory elements in those locations. In situations where two relatively small exons were close enough to each other, a single amplicon was designed to cover them both. As the Wave DHPLC system imposes limitations on the maximum length of an amplicon, comparatively large exons were covered by two overlapping amplicons.

A pair of oligonucleotides to be used as PCR primers was designed for each amplicon (see Table 3.2), using the web interface of the Primer3 program at the Whitehead Institute for Biomedical Research website and requested from commercial suppliers (Sigma-Genosys, Invitrogen).

The melting temperature (Tm) for each primer was calculated using the formula:

$$Tm=63.72+0.41(\%GC)-600/n$$

where %GC is the GC content of the oligonucleotide expressed in percentage and n is the length of the primer in base pairs. If the difference of Tm of the primers of each pair was smaller than 3°C each pair the annealing temperature (Ta) was set 2 to 3°C above their mean Tm. When the difference exceeded 3°C a "touchdown" PCR protocol was employed

by performing the annealing of the first cycle 8°C above the mean Tm and reducing the Ta by 0.5°C over the next 15 cycles. The remaining 15 cycles were performed using the Ta calculated as above.

PCR was carried out first using control DNA and 2% agarose gel electrophoresis to test the amplification conditions. If a single, strong band of the expected size was observed, 50µl PCR reactions were then carried out in patient DNA samples.

A final hybridization step was carried out, starting at 95°C and reducing the temperature 1.5°C per minute to 25°C in order to favour the formation of heteroduplexes.

The sequence of each individual amplicon was analyzed using the Navigator software (Transgenomic) to determine the melting profile of each DNA fragment and the select the temperatures at which its different melting domains (parts of the molecule that show approximately the same secondary structure at specific conditions) are predicted to adopt a percentage of helical forms between 40% and 90%.

Table 3.2 Primers used for *MYH6* mutational analysis

Primer	Sequence	Primer	Sequence
MYH6 Promo-F	aaaggagaggctggggaac	MYH6 E20-R	tagtgcatgcctccctttc
MYH6 Promo-R	catcccaccccaaacctc	MYH6 E21-1F	ccatgattgggaagctctct
MYH6 E1-F	cagagccaaaggatcaaagg	MYH6 E21-2R	taatctaggggaggggagga
MYH6 E1-R	gcaggagactcagaatgatgc	MYH6 E22-F	ccagggactgggagtctagg
MYH6 E2-F	tctgactccctggtctgtcc	MYH6 E22-R	ctgggagtcttgaggagacc
MYH6 E2-R	ctggagtatgctaagggttgg	MYH6 E23-F	caggctggtgatctttgacc
MYH6 E3-F	cagaggacaaagccactcg	MYH6 E23-R	aaatcctgcaagcacaaagg
MYH6 E3-R	tttctccagccctctcagc	MYH6 E24-F	tttagaaggaggcaaaagagc
MYH6 E4-F	ctgggaggaggtcagtgg	MYH6 E24-R	gcaccctgcactctatctacc
MYH6 E4-R	ccccctggcttatttagg	MYH6 E25-F	agagaatgagccccagagg
MYH6 E5-6F	agctgcaggaggagtagagc	MYH6 E25-R	ccagatattgtgtagaaccctaagc
MYH6 E5-6R	tcccagccttaaacctctcc	MYH6 E26-1F	accactgcttgagaggaacc
MYH6 E7-F	gatgctgagccctgtatgg	MYH6 E26-1R	tctccttctccagcttctgc
MYH6 E7-R	ggagggttaggggtaactcg	MYH6 E26-2F	gccgagttccagaagatgc
MYH6 E8-F	ggtaggatcctgtggagtcg	MYH6 E26-2R	tggcagacagagagagaagg
MYH6 E8-R	ccaaagcctatgctctcttcc	MYH6 E27-28-F	ccctctcttccttcctctgg
MYH6 E9-F	ttctgggctgaacagagg	MYH6 E27-28-R	ctggcactgagatgaattgc
MYH6 E9-R	ggcaggaatgatgagactgg	MYH6 E29-F	tctagagaatggggcacagg
MYH6 E10-F	catggccacctttttctgg	MYH6 E29-R	tccacttccgtctcatgacc
MYH6 E10-R	gcatgcaggagtcgttgg	MYH6 E30-F	ccagtagagtcacacacacacc
MYH6 E11-12-F	caactctacctgccccttcc	MYH6 E30-R	ctttggcctctcactgaacc
MYH6 E11-12-R	atctgagtcccgcagagagc	MYH6 E31-32-F	agggctggggagctaagg
MYH6 E13-F	caagcctgggtgacagagc	MYH6 E31-32-R	agacagcggcagaacagg
MYH6 E13-R	caagcgagtgattgttctcc	MYH6 E33-F	cgatagtcctggctgacacc
MYH6 E14-F	ggaggggacagccatacc	MYH6 E33-R	ccagacaccactgcttctcc
MYH6 E14-R	gggtgtagaagggactcagc	MYH6 E34-F	ggagaagcagtggtgtctgg
MYH6 E15-F	gtcagggtatgggactgtgg	MYH6 E34-R	ctagatgtcctgggctctgc
MYH6 E15-R	ctgcctatggagtcatgtgc	MYH6 E35-F	ggagaaagggtatgaaatcagg
MYH6 E16-F	gggctcctttattttccagc	MYH6 E35-R	gccttgtttctgtctttaggg
MYH6 E16-R	ataggtggtgcagccagaag	MYH6 E37-F	gaggagggaaaggtgattgc
MYH6 E17-18-F	gagtgtctgggacagggttt	MYH6 E37-R	gtccaggcccctctgtagg
MYH6 E17-18-R	atctgctctgcccacagaat	MYH6 E38-F	aggctgagaatcccatagcc
MYH6 E19-F	tctgtgggcagagcagatc	MYH6 E38-R	gaagggcacccatatcagg
MYH6 E19-R	aagccagaattaggcttctgc	MYH6 E39-F	aagggcatctcccttaggc
MYH6 E20-F	accctggatactcccctctg	MYH6 E39-R	tctggcagctctgatacagg

A PCR reaction was performed per patient per amplicon (more than eleven thousand in total). Using patient DNA as template, PCR reactions were performed, heteroduplexes formed and injected in the dHPLC column using the Wave system at the appropriate temperatures.

3.5 Results

A summary of coding variations found in this study is represented in Figure 3.1. Within the 494bp amplicon corresponding to a segment of the *MYH6* promoter only one variant was found. It was observed in patient with patent foramen ovale and Ebstein's malformation and consisted in a transition G→A 370 bases upstream the transcriptional start. It is not localized in an already known transcription factor binding site but it lies in the middle of a 14bp segment that has been conserved through evolution. DNA samples from two brothers of the patient were screened for the variant, but they were found to be non-carriers (Figure 3.2).

Two variants were found in the part of the gene that encodes Loop-1, a segment of the myosin head involved in ADP release (Murphy and Spudich 1998). 1) D208N (aspartate, a charged polar residue for asparagine, uncharged polar) was found in 9 patients with different CHDs and in two control DNA samples) (Figure 3.3.) and 2) The N211S variant, that replaces asparagine (uncharged polar) for serine (same group). It was observed in a male patient with aortic valve stenosis for bicuspid aortic valve (Figure 3.4).

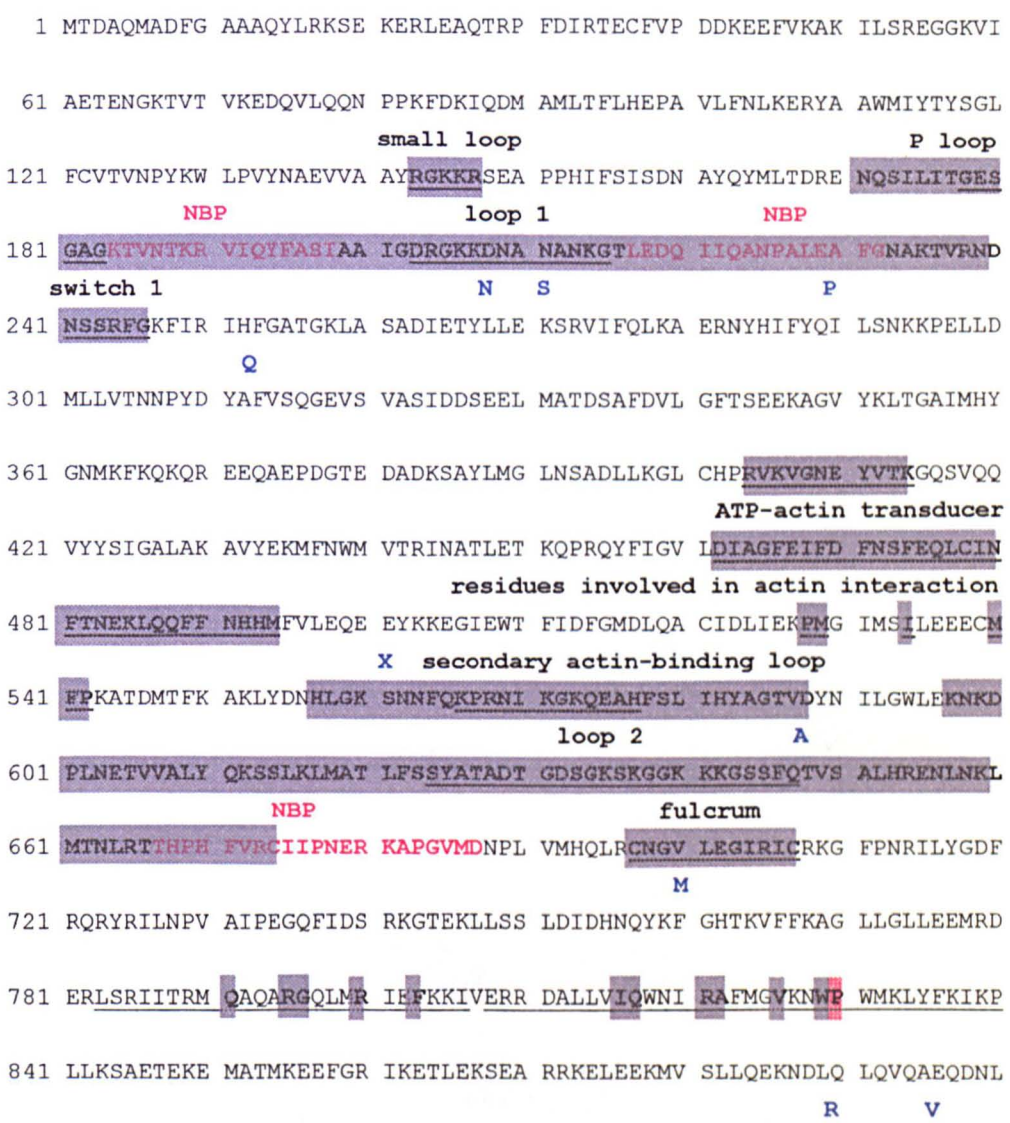


Figure 3.1: Diagramatic representation of the myosin head. The grey highlighted sequence represents conserved segments of the molecule. The nucleotide binding pocket regions are indicated by red letters. The blue letters indicate mutations found in our cohort (see text).

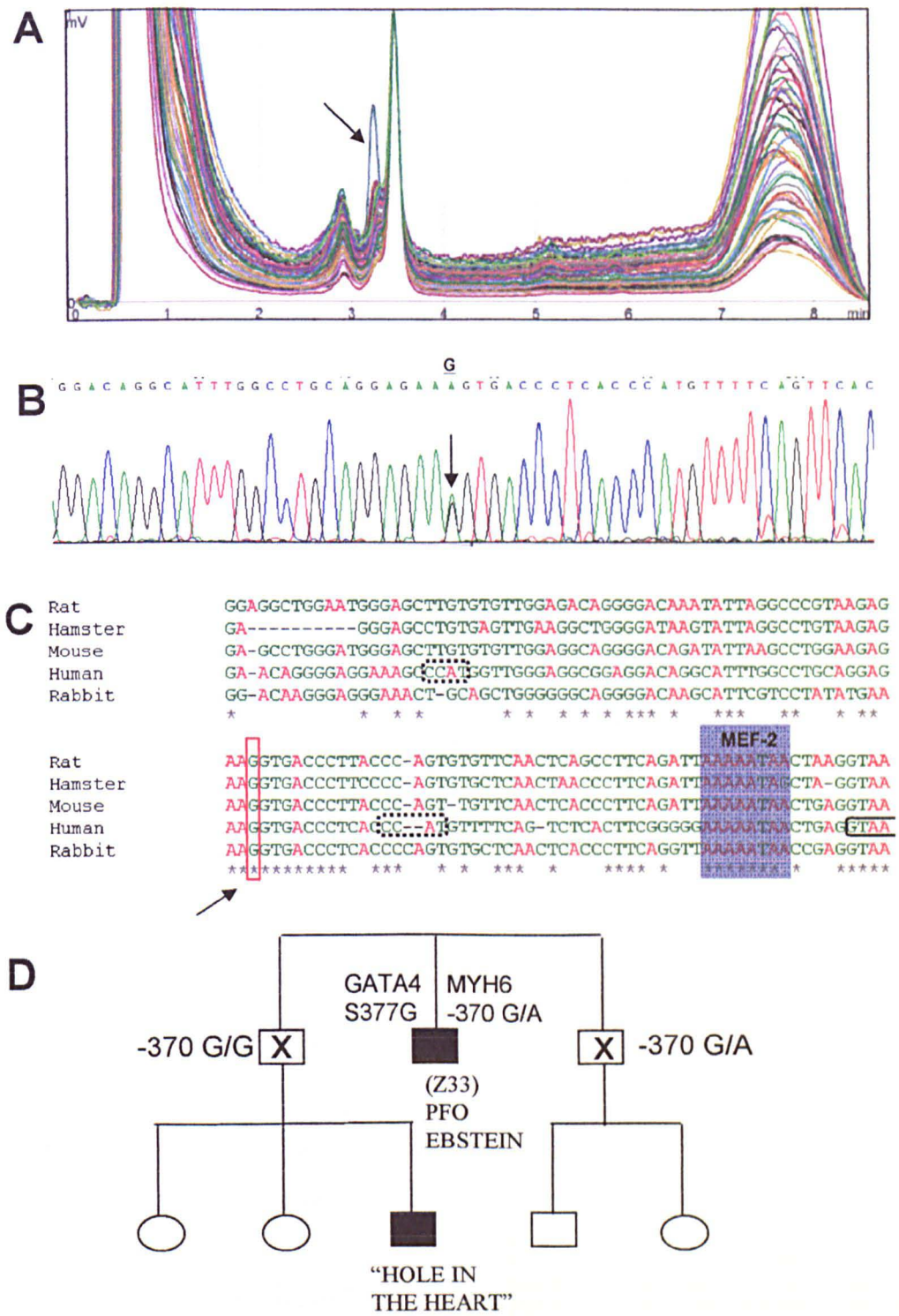


Figure 3.2: Sequence variant within the promoter amplicon of MYH6. (A) dHPLC trace difference in a single sample. The sequence (B) showed a transition G→A in the middle of a 14bp conserved segment (C). Two brothers of the patient were screened; one of them has a son with an identified heart defect (D), no variant found in them.

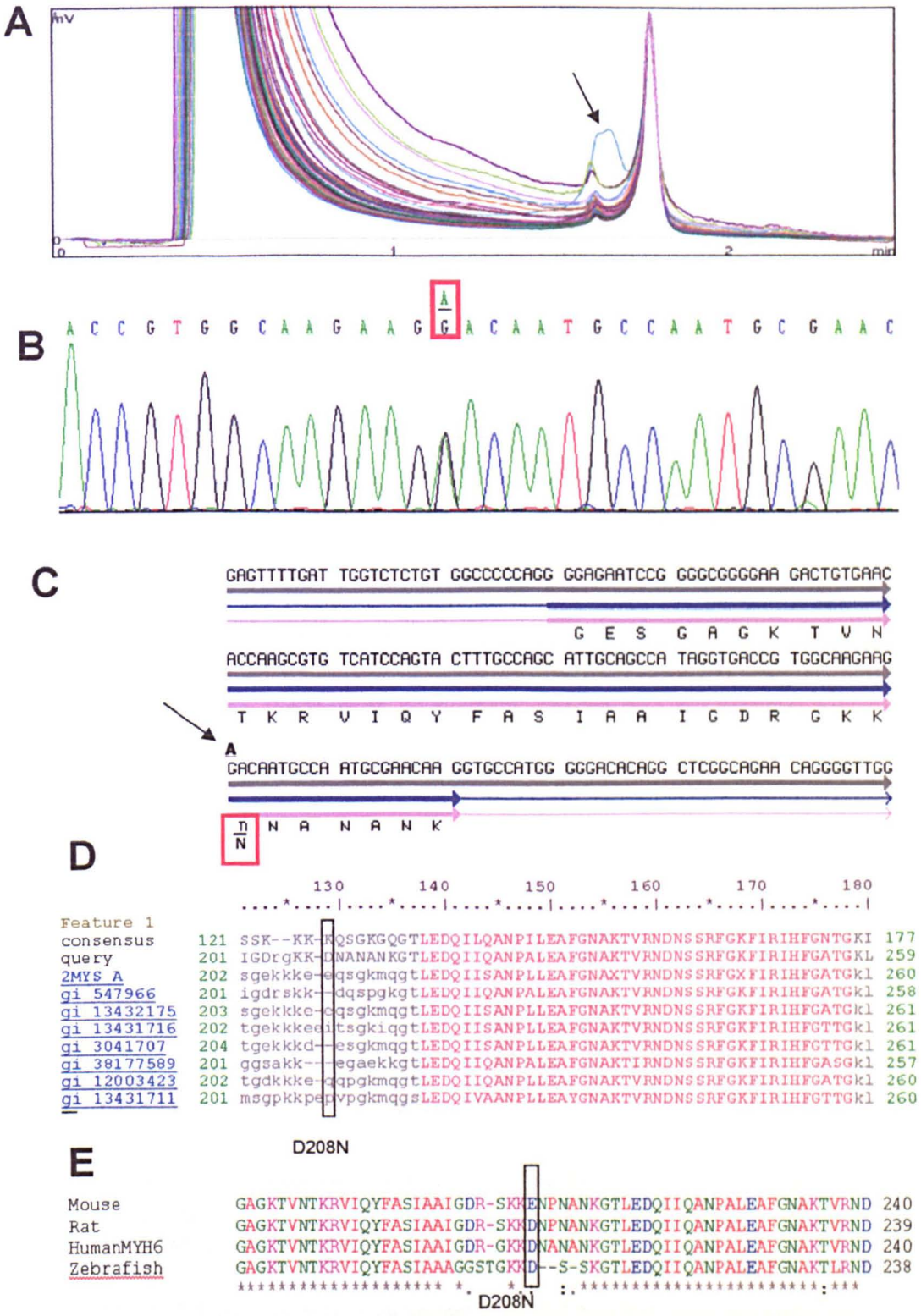


Figure 3.3: A sequence variant in Exon 7 (A), consisted of a transition G→A (B), predicted to produce a D208N (C) protein variant in a residue not conserved in an "all myosins" context (D) but conserved when compared with the cardiac myosins of other species (E).

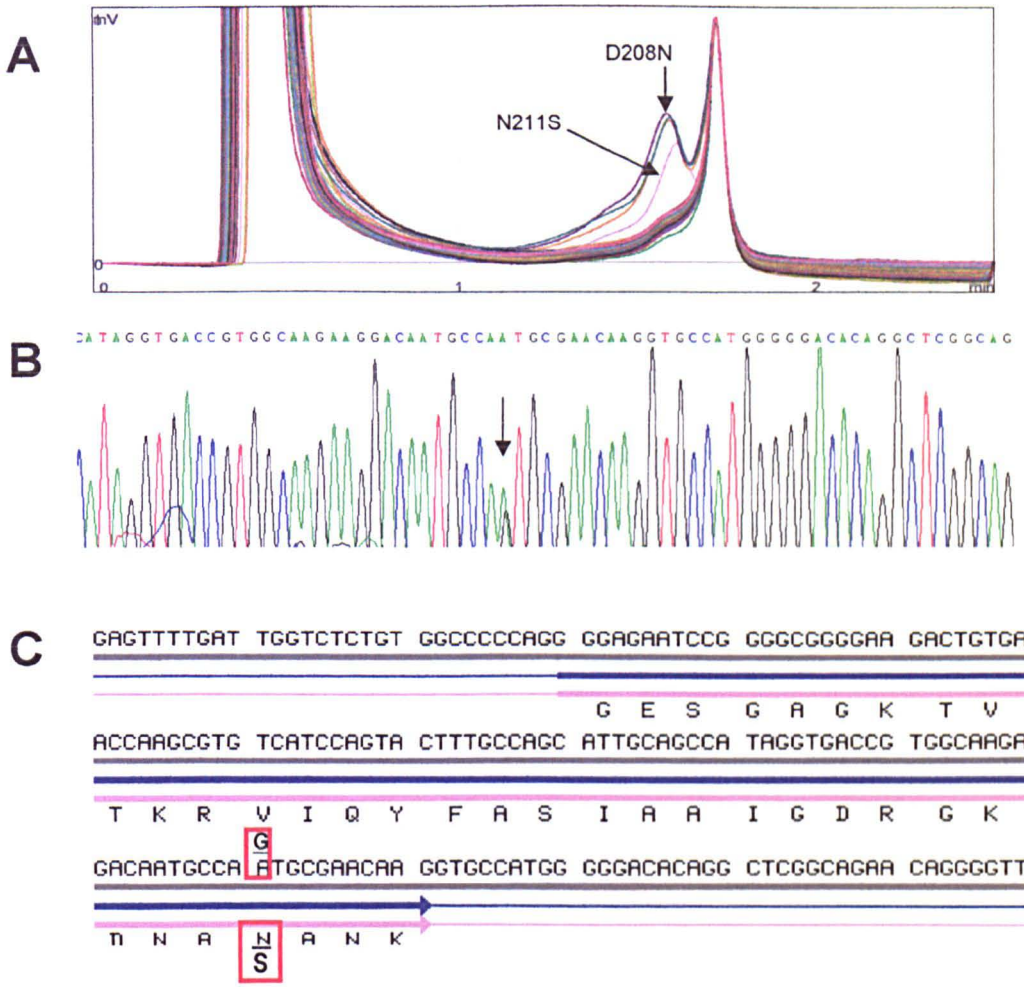


Figure 3.4: Another sequence variant in Exon 7 besides other two instances of the D208N variant (A), consisted of a transition A→G (B), predicted to cause a N211S (C) protein variant in a residue not conserved in an "all myosins" context (D) but conserved when compared with the cardiac myosins of other species (E).

A230P (alanine for proline, both nonpolar) was found in a patient with atrial septal defect. A Prof prediction of secondary structure locates this position in a potential α-helix portion of the nucleotide binding pocket (Figure 3.5). Due to its lack of a hydrogen atom on the α-amino group, when a proline residue is located in an α-helical segment of a protein, it induces an angle in the axis of the helix (Figure3.6). This residue is highly conserved in many myosins of many species and is located in the nucleotide binding pocket of the molecule.

H252Q (histidine, charged polar; glutamine uncharged polar) was identified in a patient with transposition of great arteries. The healthy mother of the patient carries the variant. This residue is not located in a conserved segment of the protein (Figure 3.7).

A nonsense mutation (E501stop) was found in a patient with tricuspid atresia, ventricular septal defect and hypoplastic right ventricle. The transcript encoded by the mutated copy of the gene would produce, if translated, a truncated protein of 500 residues. However, it seems more likely to undergo nonsense mediated decay instead. The same mutation was found in the mother and maternal grandfather of the patient, neither of whom is affected (Figure 3.8).

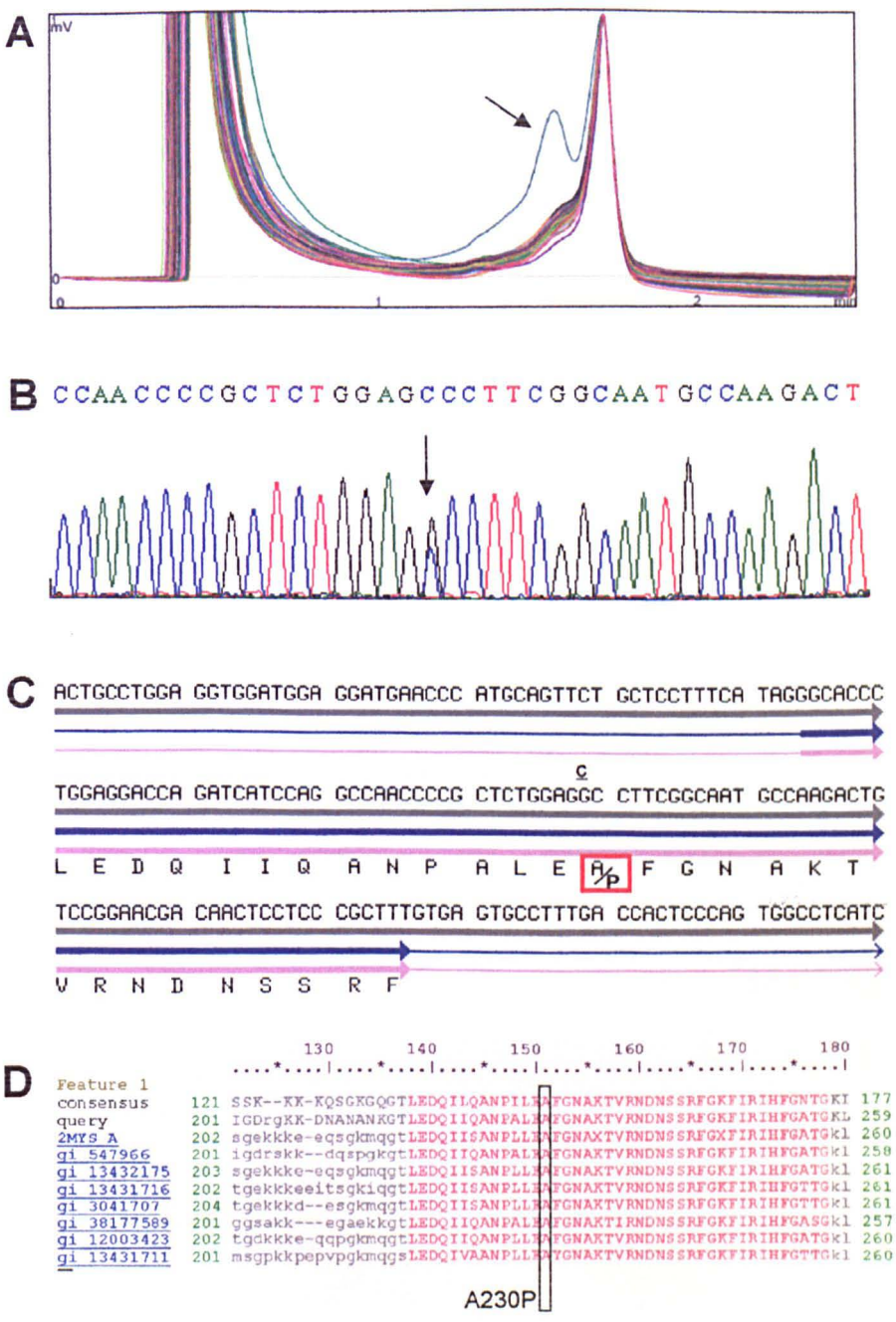


Figure 3.5: A variant in exon 8 (A), consisted of a transversion G→C (B), predicted to cause a A230P (C) protein. This residue is highly conserved in many myosin molecules (D).

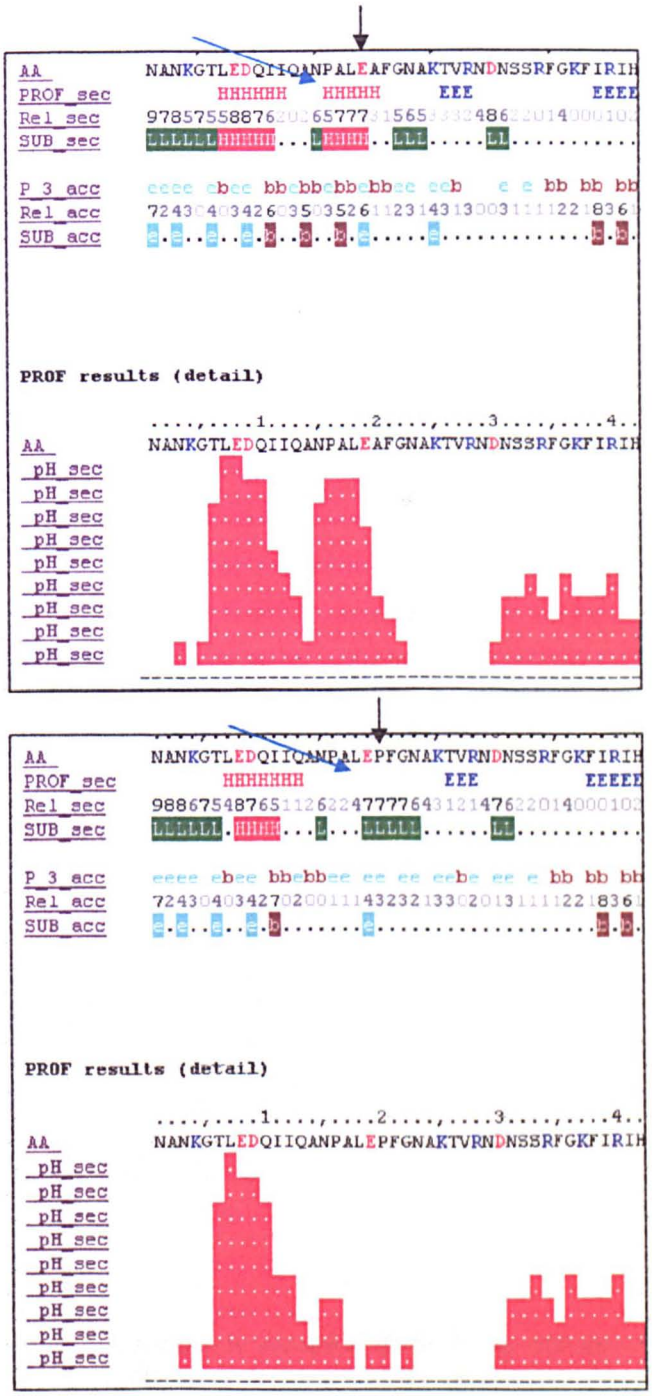


Figure 3.6: A Prof prediction of secondary structure of a segment of MYH6 processed using the sequence of the wild-type (upper panel) protein and the A230P variant (lower panel). Black arrows point to the relevant residue. According to the prediction, the segment of the variant protein containing residue 230 fails to adopt a helical conformation present in the wild-type (blue arrows).

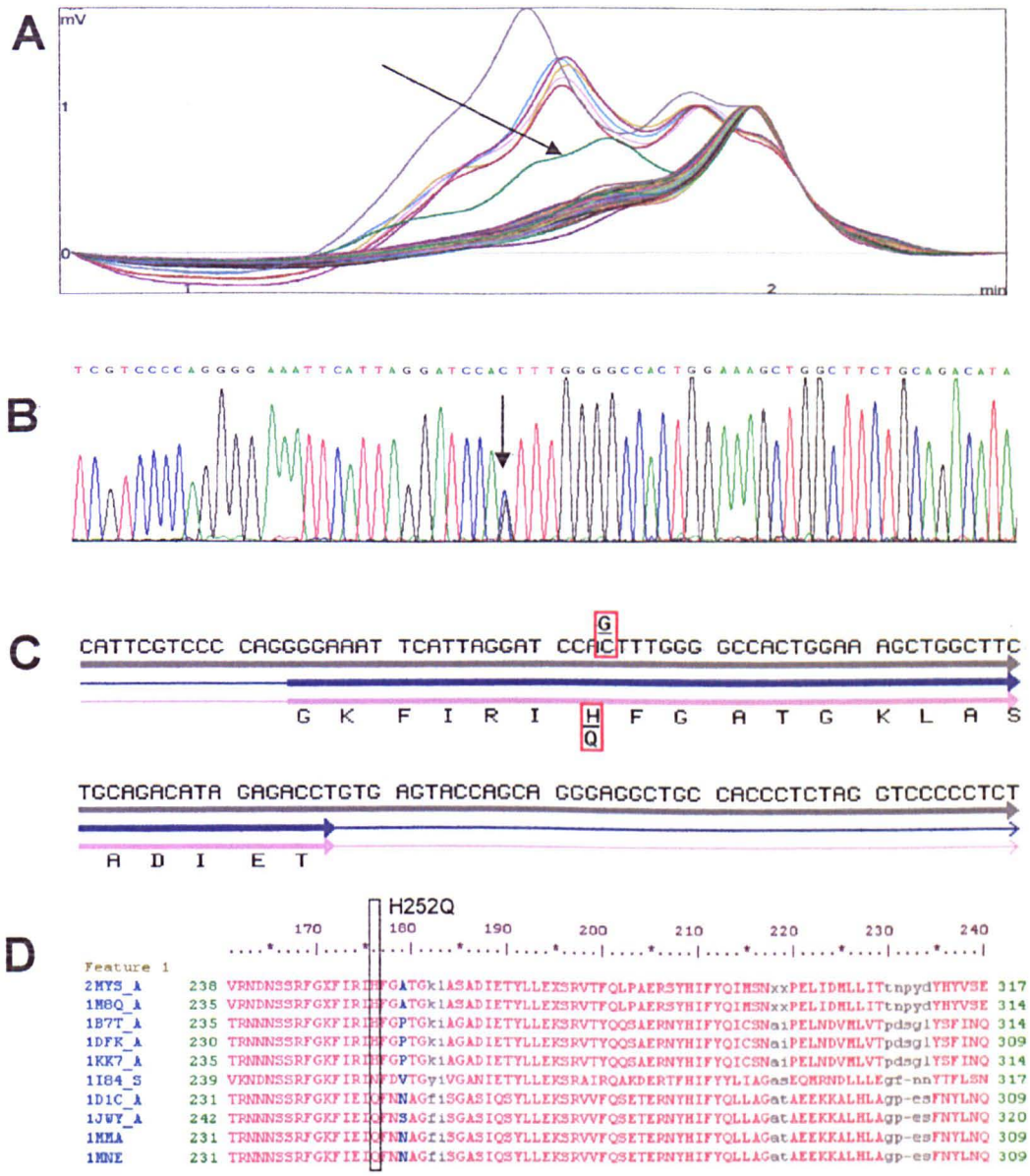


Figure 3.7: A sequence variant in Exon 9 (A), consisted of a transversion C→G (B), predicted to cause a H252Q variant (C). This residue is not in a conserved position (D), unlike the flanking residues.

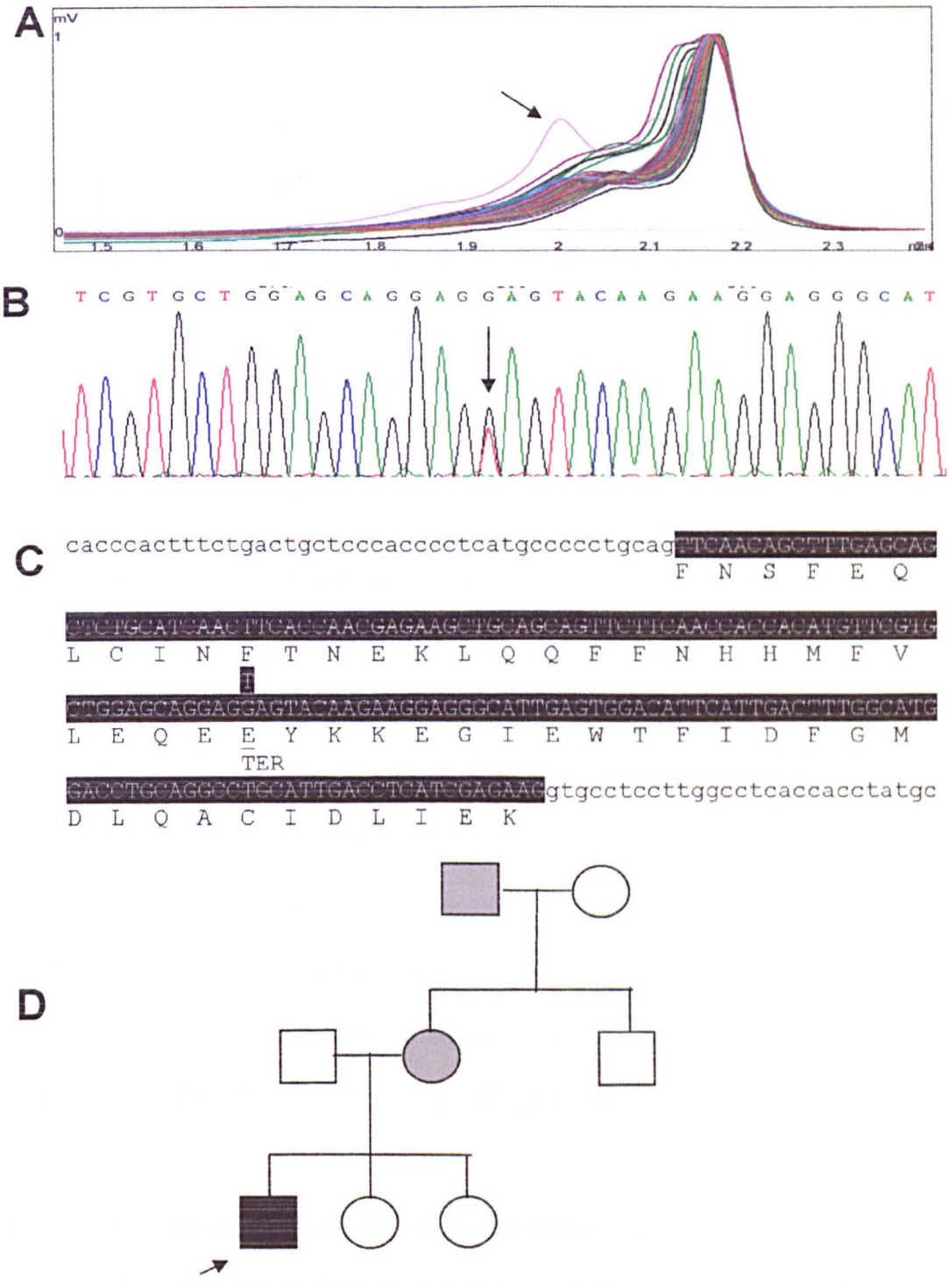


Figure 3.8: A sequence variant in exon 14 (A) consisted in a transversion G→T (B) that predicted an aberrant stop codon at position 501 (E501Stop). The mother and maternal grandfather are healthy carriers of the mutation.

A one base pair deletion (Figure 3.9) was found in 12 (various phenotypes) of the 380 patients, 31 base pairs upstream the 5' end of exon 14, in the putative splicing branch site, modifying its sequence from ttctgac to ttcgac (consensus [c/t]n[c/t]t[a/g]a[c/t] or ynytray IUPAC code). This variant has been found in four of 192 control subjects screened.

D588A (aspartate, charged polar for alanine, non-polar) was discovered in a patient with atrioventricular septal defect. This residue is located in the C-terminal portion of a sequence that harbours the secondary actin binding loop. This particular position is non-conserved (Figure 3.10)

V700M (valine for methionine, both non-polar) was found in a patient with a large foramen ovale. Two healthy sons of the patient were screened and found negative for the variant. This valine residue is highly conserved across many myosins from many species and is located in the middle of the fulcrum region of the myosin head (Figure 3.11).

The A895V (alanine for valine, both non-polar) variant was found in a patient with aortic coarctation and occurs in a non-conserved residue of the region of the molecule between the neck and the tail region (Figure 3.12).

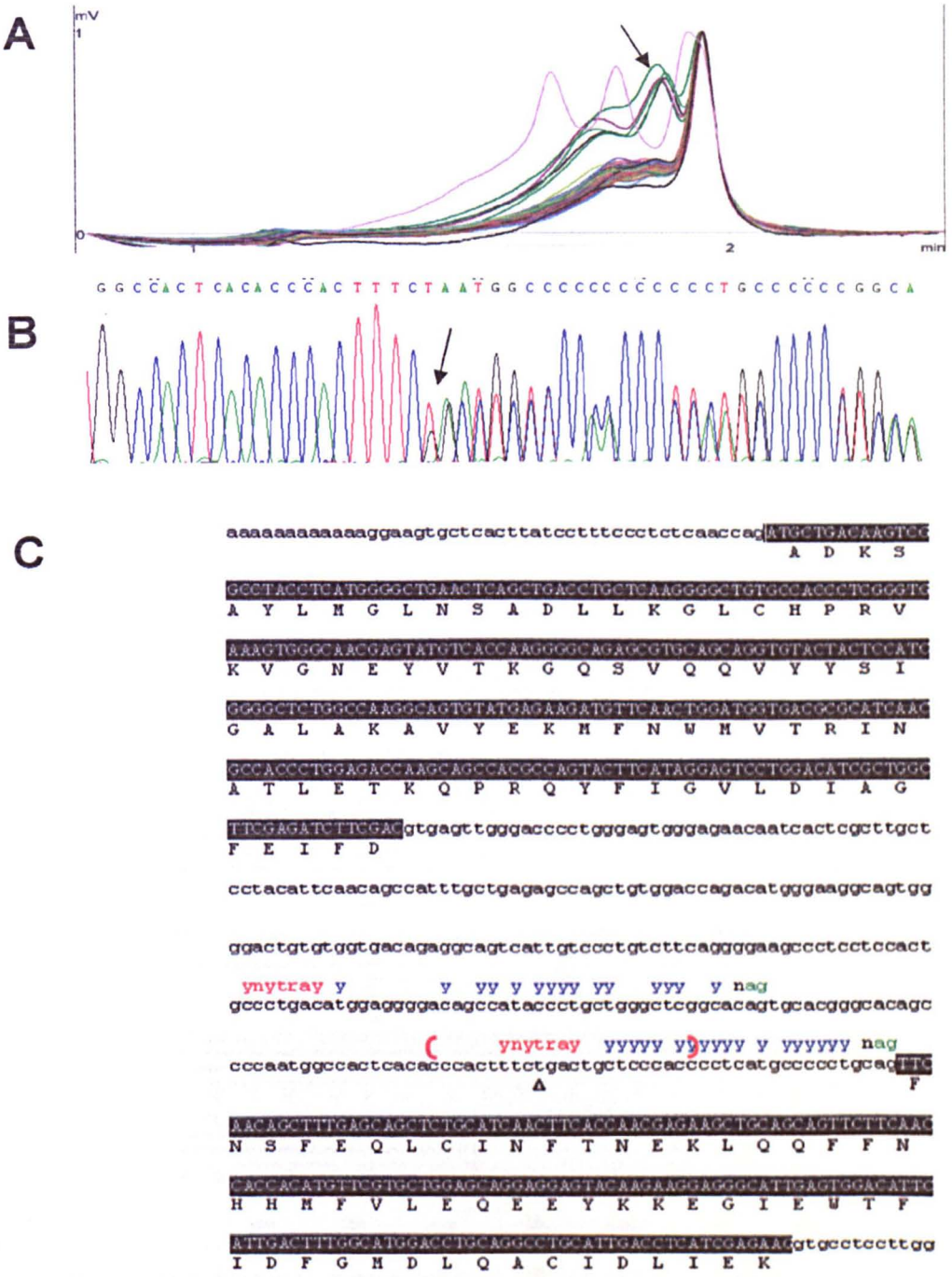


Figure 3.9: A sequence variant in amplicon for exon 14 found in 12 patients (A), consisted of a deletion (B) of a thymine nucleotide 31bp upstream the 5' of exon 14, modifying the putative branch site (arrow) of the intron (C). The altered branch lies "within range" (inferior "ynytray" between red parenthesis) of the acceptor splicing site of exon 14 (inferior green "ag"). The only potential alternative branch site in the intron (superior ynytray), if used, would use a cryptic acceptor site (superior green ag), and induce a partial intron retention in the mature transcript.

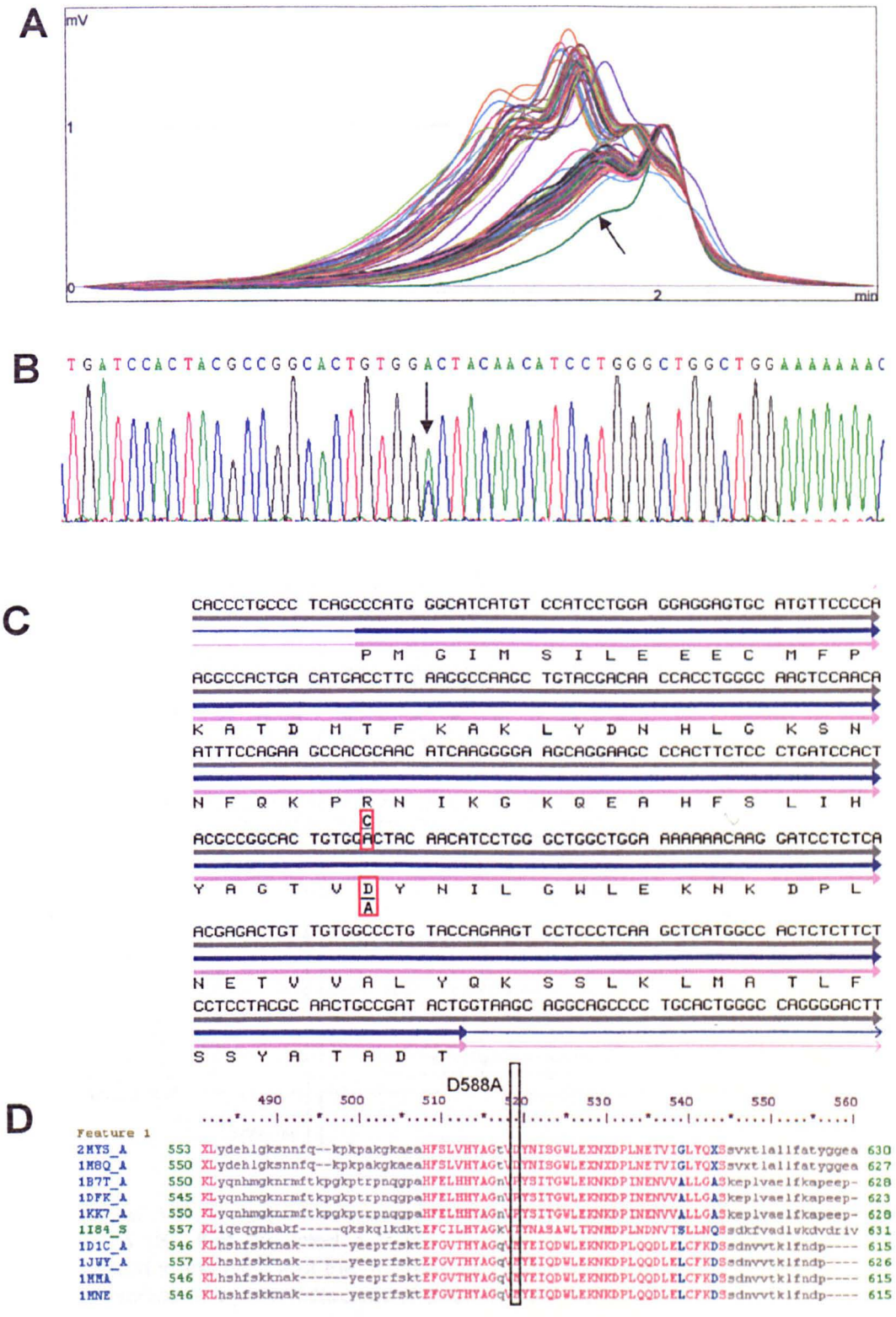


Figure 3.10: A sequence variant in exon 15 (A) consisted in a transversion A→C (B), that predicted a D588A variant protein (C). Even though this is a well conserved segment of the protein, the relevant residue itself is not (D).

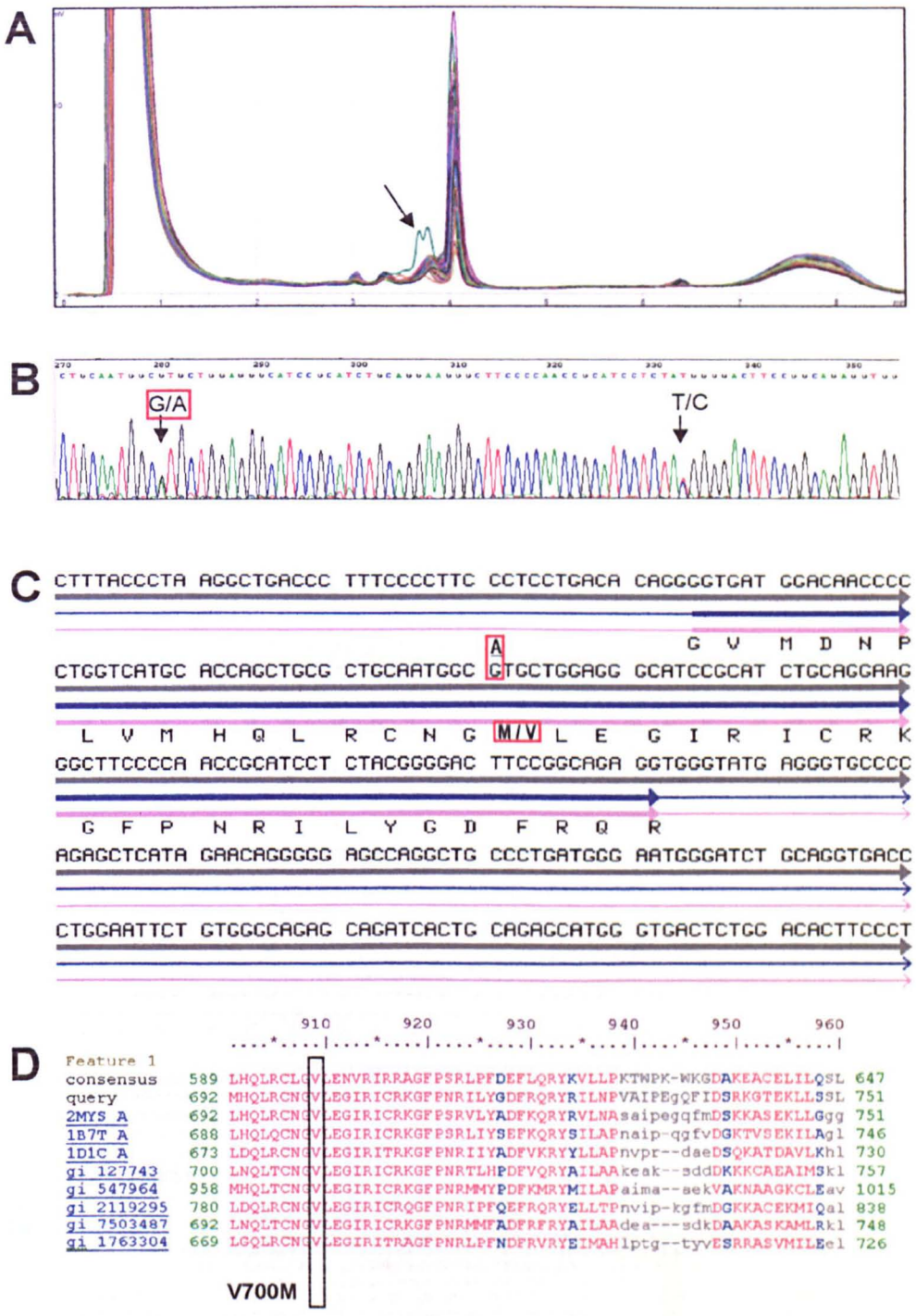


Figure 3.11: A sequence variant in exon 18 (A) consisted in a transition G→A (B), that predicted a V700M variant protein (C). This residue is located in the middle of the "fulcrum" of the myosin head and is extremely conserved in myosin molecules of diverse species (D).

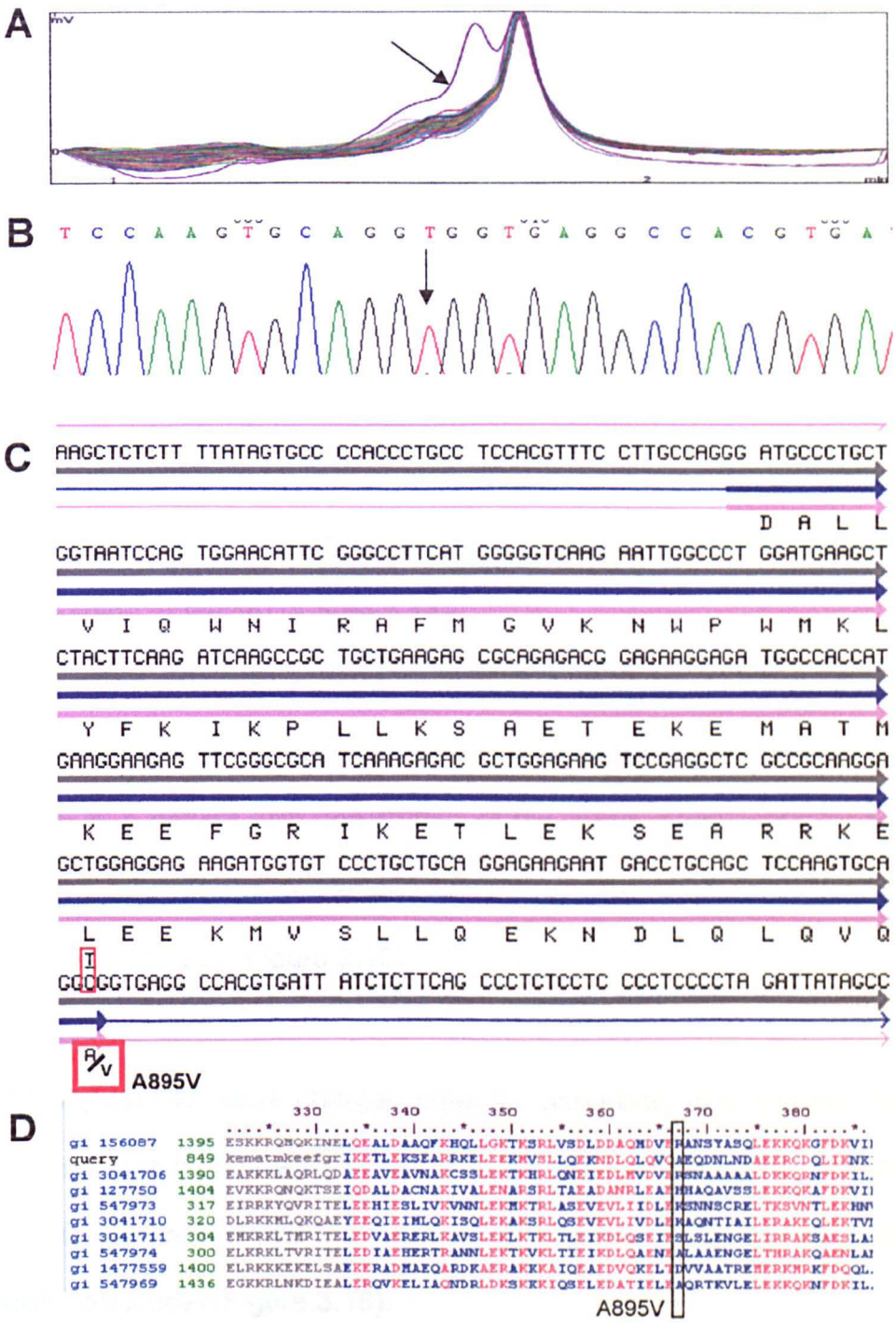


Figure 3.12: A sequence variant in exon 21 (A) that consisted in a transition (B) C→T (as the primers used for amplification did not yield a readable sequence, it was necessary to clone the PCR product, here just the segment from the variant chromosome is shown). The variant is predicted to encode a A895V protein. The residue in that position does not show conservation.

R111S (arginine, basic charged polar for serine, uncharged polar) was detected in a patient with atrial septal defect. This residue is conserved in many myosins of many species and (as all variants described below) is located in the myosin tail (Figure 3.13).

A1366D (alanine, nonpolar for aspartate, charged polar) was discovered in a patient with aortic stenosis and dysplastic aortic valve. This position is occupied in other myosins by non polar aminoacids (alanine, phenylalanine or valine) (Figure 3.14).

T1379M (threonine, uncharged polar for methionine, nonpolar) was detected in a patient with coarctation of the aorta, atrioventricular septal defect and multiple ventricular septal defects. The variant is not located in a conserved residue (Figure 3.15).

R1422Q (arginine, basic charged polar for aspartate, acid charged polar) was found in a patient with atrioventricular septal defect. This position is occupied in other myosin by lysine or arginine (both basic charged nonpolar residues) (Figure 3.16).

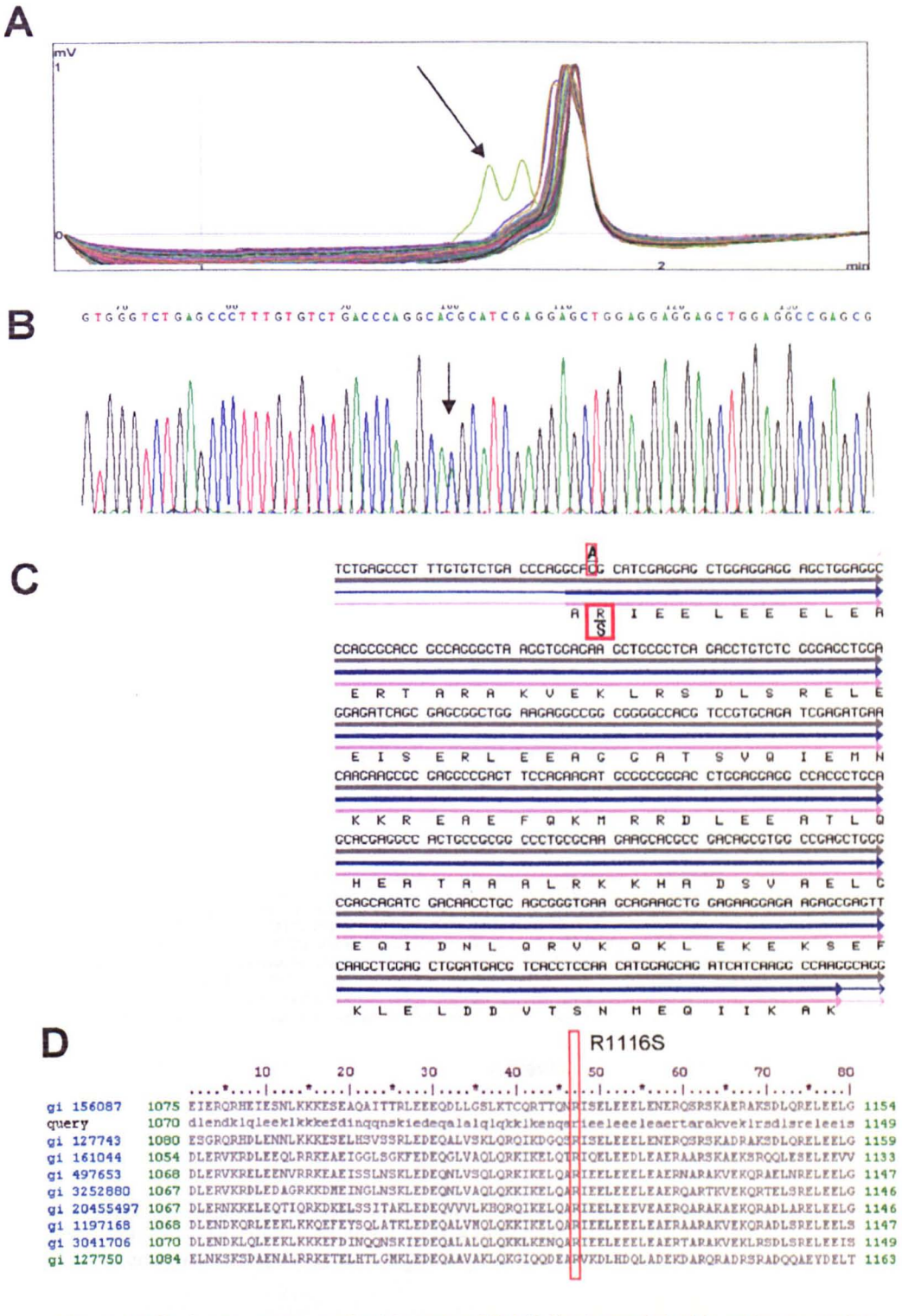


Figure 3.13: A sequence variant in exon 26 (A) that consisted in a transversion C→A (B), that predicted a R1116S variant protein (C). This residue is located in the tail or the myosin molecule and is highly conserved in various of many species (D).

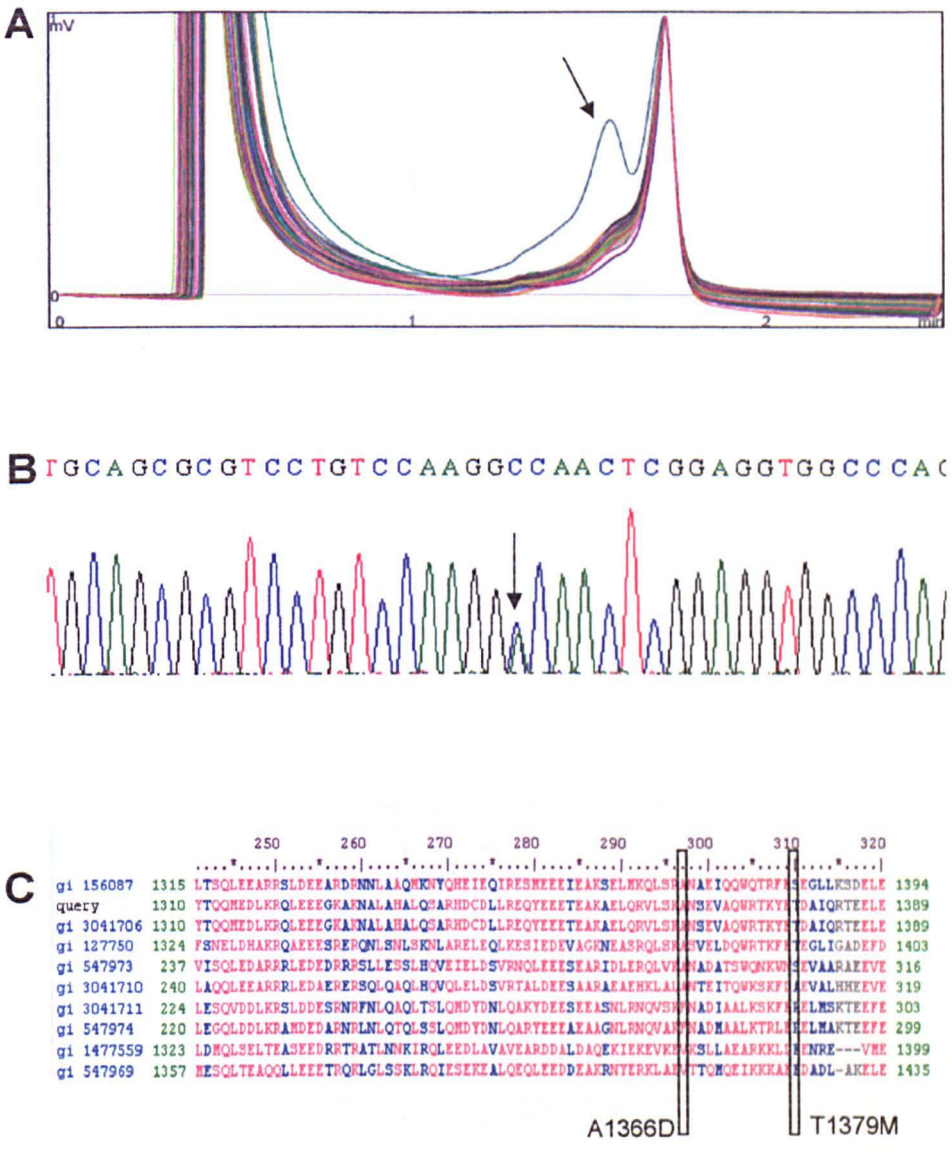


Figure 3.14: A sequence variant in the exon 29 (A) that consisted in a transversion C→A (B), that predicted a variant protein A1366D (Figure 3.15C). That position is not conserved in other myosins (C).

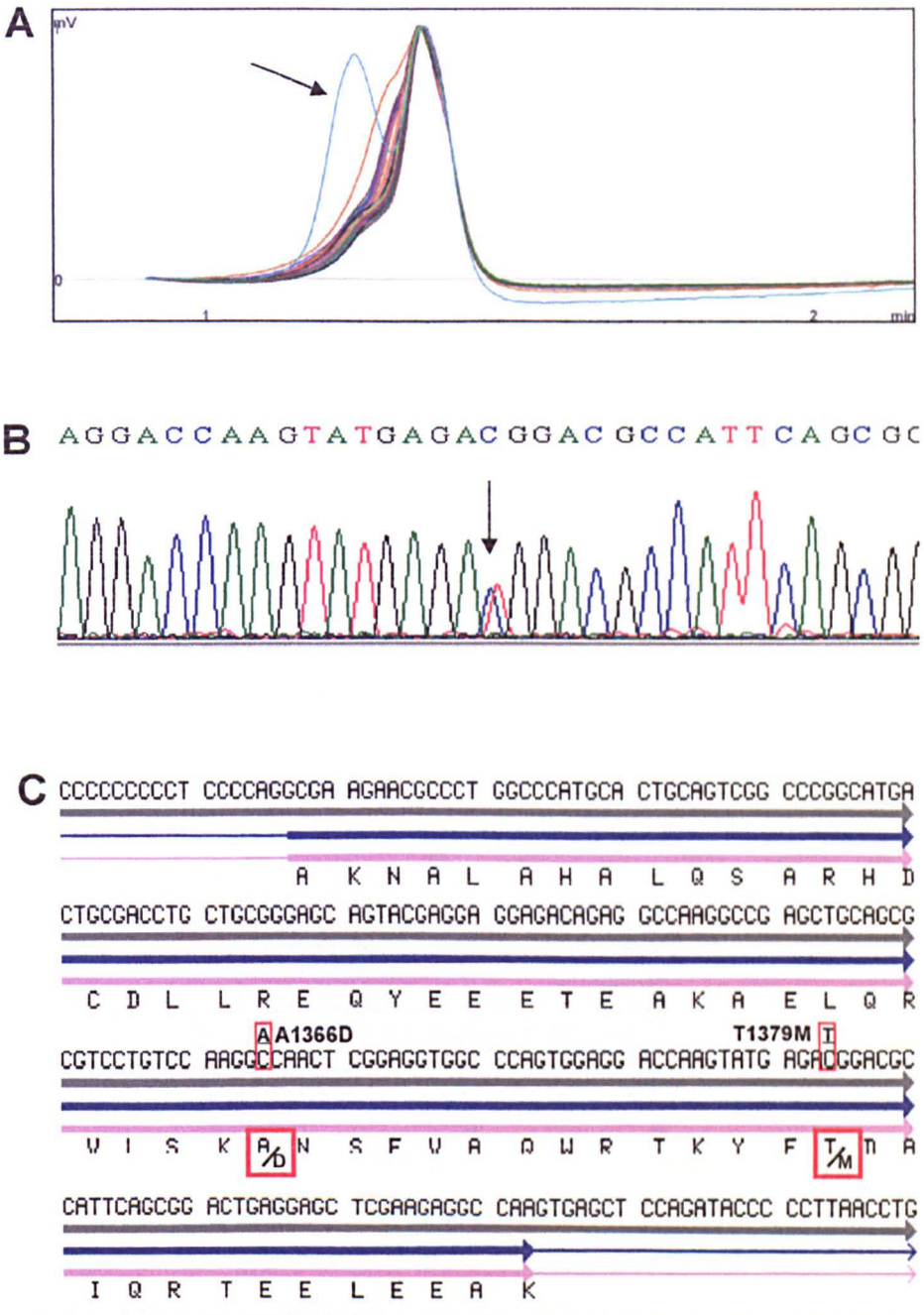


Figure 3.15: A sequence variant in the exon 29 (A) that consisted in a transition C→T (B), that predicted a variant protein T1379M. That position is normally occupied in other myosins by alanine, valine or phenylalanine, all three nonpolar residues (Figure 3.14C).

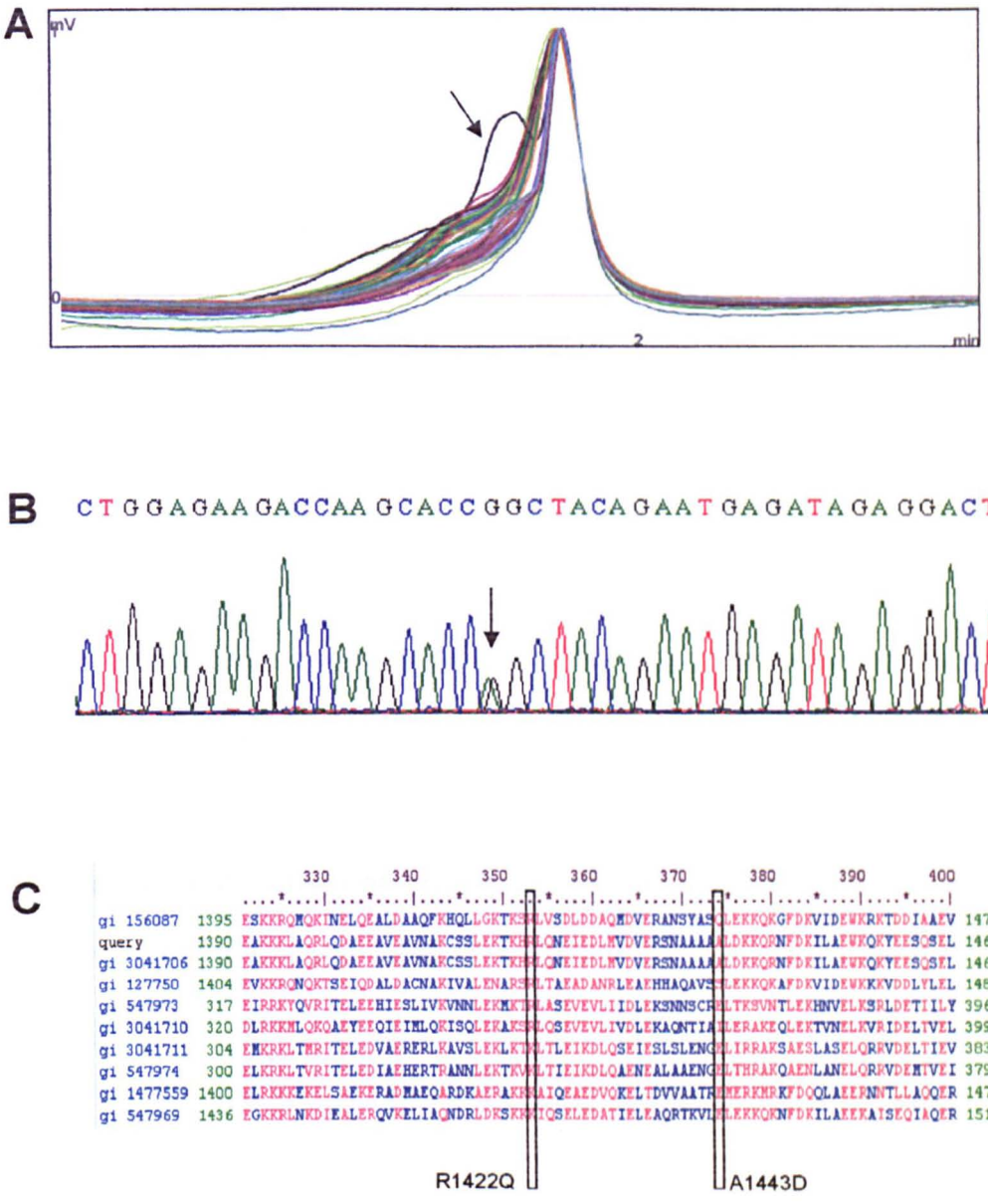


Figure 3.16: A sequence variant in the exon 30 (A) that consisted in a transition C→A (B), that predicted a variant protein A1422Q (Figure 3.17C). The residue in that position is occupied by arginine or lysine in other myosins (C).

A1443D (alanine, non-polar for aspartate, charged polar) was discovered in a patient with atrial septal defect in addition to a transition G→A located 50 base pairs downstream the 3' end of exon 30 (both inherited from the mother) and a transition T→C 13 base pairs downstream the 3' end of the same exon (inherited from the father) (Figure 3.17).

R1865Q (arginine, charged polar for glutamine, uncharged polar) was found in a patient with atrial septal defect. This position does not show conservation in other myosins (Figure 3.18).

A mutation of the acceptor splicing site (AG→GG) at the 5' end of exon 38 was found in a patient with truncus arteriosus (Figure 3.19).

Several synonymous and non-coding variants were found within the *MYH6* amplicons. A summary can be found in Table 3.3.

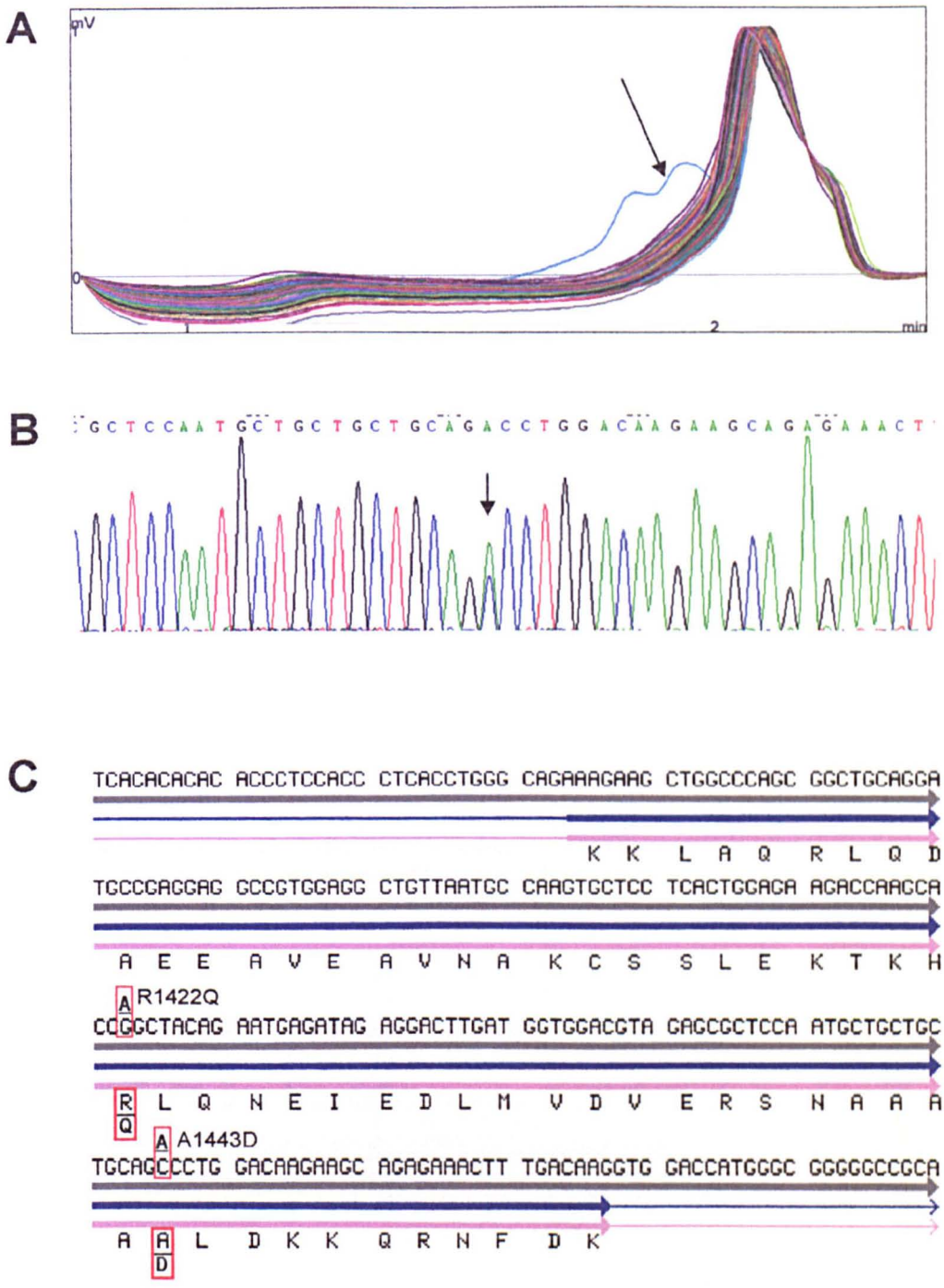


Figure 3.17: A sequence variant in the exon 30 (A) that consisted in a transversion C→A (B), that predicted a variant protein A1443D (C). The residue in that position does not show conservation in other myosins (alignment in Figure 3.16C).

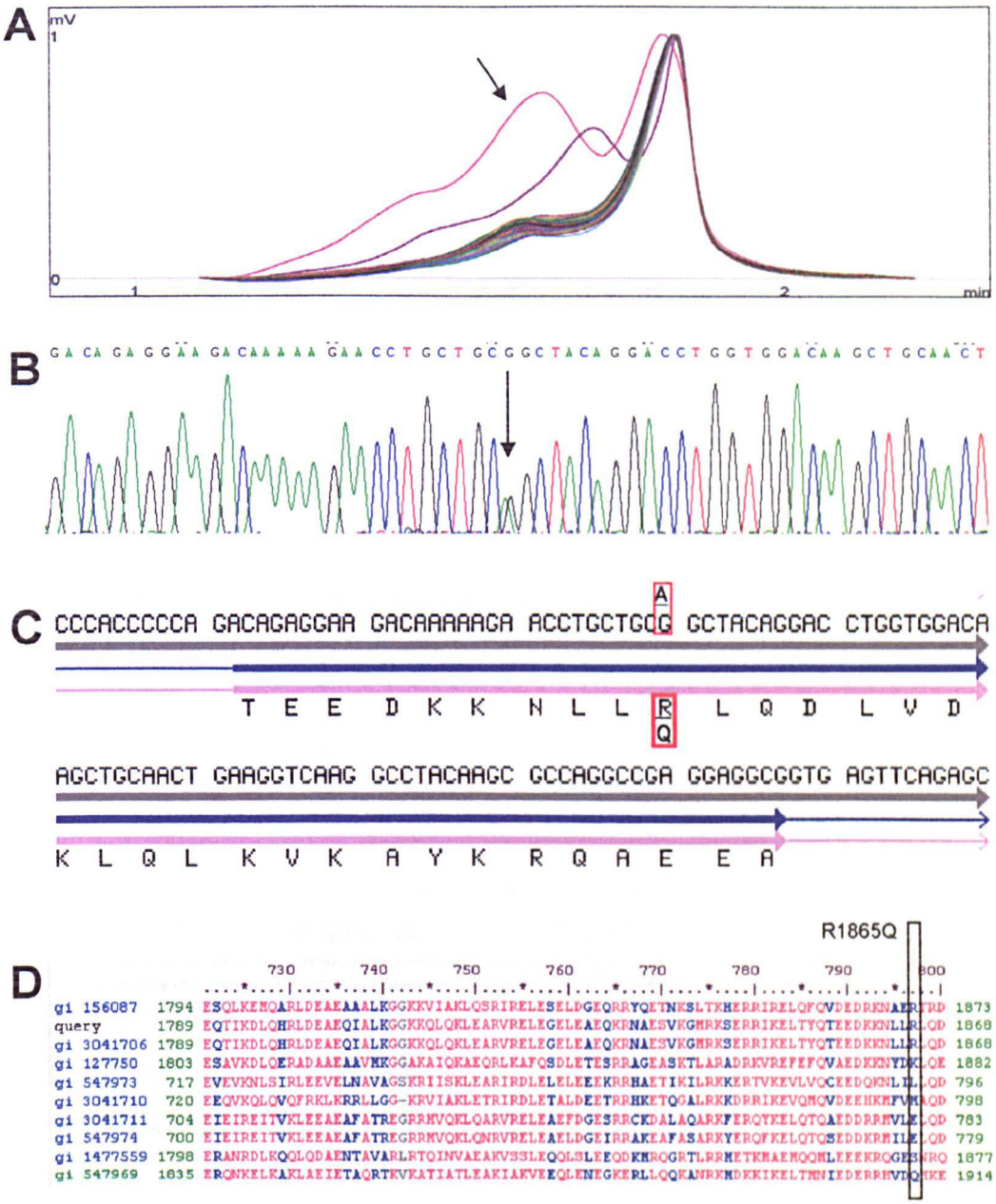


Figure 3.18: A sequence variant in exon 26 (A) that consisted in a transition G→A (B), that predicted a R1865Q variant protein (C). This residue is not conserved in other myosin molecules (D).

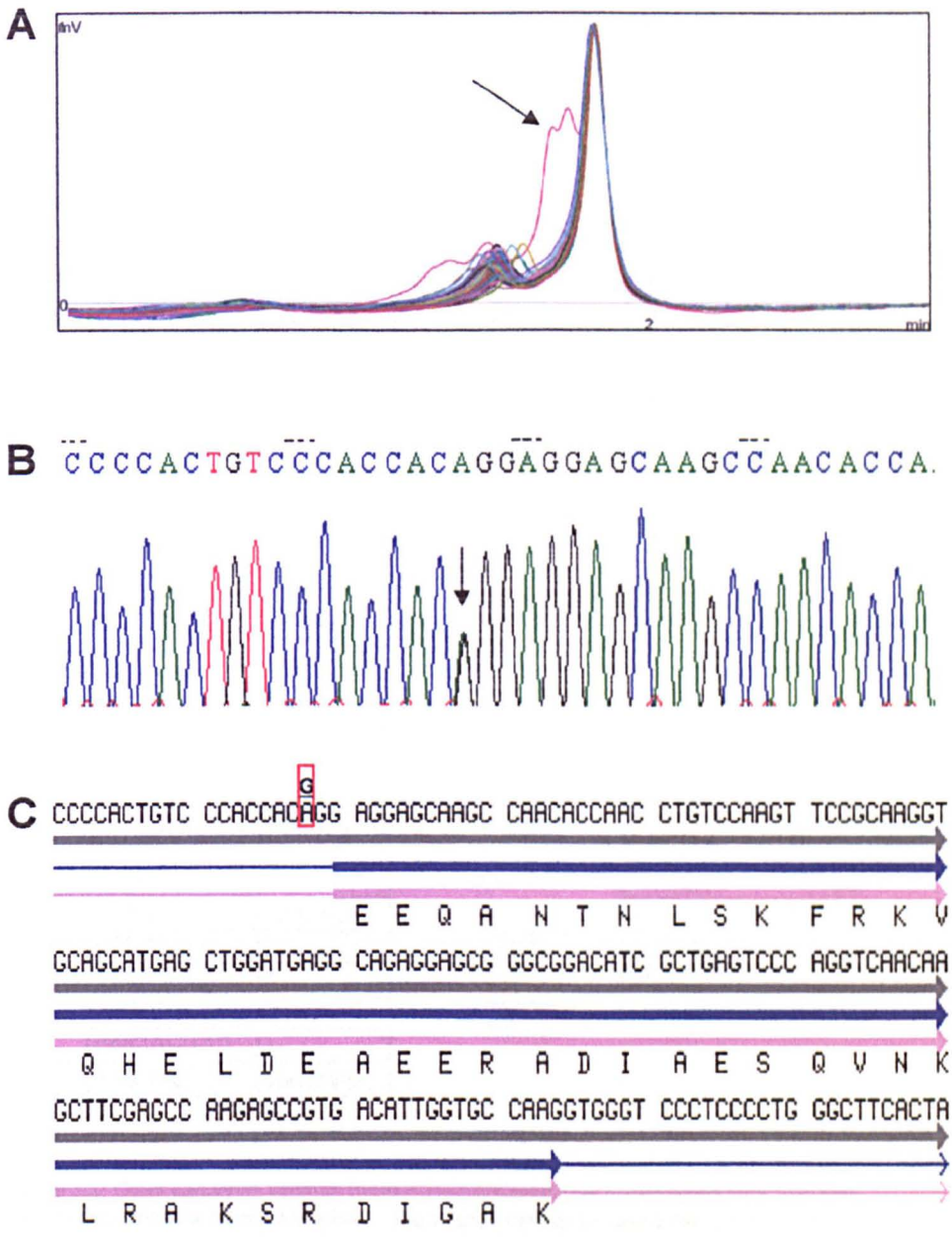


Figure 3.19: A sequence variant in the exon 38 amplicon (A) that consisted in a transition (B) A→G in the acceptor splicing site at the 5' end of exon 38 (C).

Exon	Variant	Cases
E1	+82 G-A	1
	+36 C-A	1
	+38 T-G	1
	6 G-C	1
E2	+96 C-T	3
E3	-41G-A	1
	+16 C-T	27
E4	A110A	2
E5-6	L131L	90
	E134E	90
	R143R	5
	+63 T-G	1
	+14 C-A	1
	-81C-G	2
E8	+119 G-A	1
E9	-34 G-A	12
E10	L291L	1
	-11 G-A	common
E11-12	L323L	1
	-29 C-A	2
E13	N445N	22
E14	-26 G-A	1
	-69 C-G	1
E15	-39 T-C	common
	-78 T-G	1
	E538E	1
E16	+31 C-T	common
E17-18	+15 C-A	5
	Y717Y	4
	N663N	3
	delGAG+10	1
E19	+18 T-C	common

Exon	Variant	Cases
E22	-35 T-C	4
	+4 A-G	1
E23	E981E	5
	+56 C-T	1
	+143 T-C	3
E24	L1040L	1
	G1059G	1
E26	-59 G-A	37
	-59 G-A	2
	-66 C-T	14
E27-28	-46 G-A	1
E29	S1337S	40
E30	A1443D	2
	+50 G-A	2
	+13 T-C	2
E31	delGCTAAG +97	6
	+96 G-A	2
E32	E31+71 C-T	common
	+24 C-T	12
	+42 G-A	1
	S1512S	1
	+71 C-T	3
	+74 C-T	3
E34	D1660D	common
	-86 G-A	common
	-17 A-T	4
	-9 G-A	1
E35	A1753A	common
	-22 A-G B32	common
E37	L1866L	3

Table 3.3: Other sequence variants found during the dHPLC of MYH6 amplicons. The first column indicates the relevant amplicon. The second column indicates the nature of the variation. Synonymous variants show the same one-letter code at the left and right of the aminoacid residue number (example A110A). Numbers preceded by minus (e. -41G-A) or plus (e.+82G-A) symbols indicate that the variant is located 41bp upstream the 5' end of the relevant exon or 82bp downstream the 3' end of the exon, respectively. For deletions (del) (e. delGAG+10) the sequence of the deleted segment is shown with its position. The third column represents the number of times the variant was seen in our cohort.

3.6 Discussion

dHPLC is a powerful technique for detection of mutations. The detection rates reported vary according to the source, but in general are thought to be close to 97%. However, as the procedure is based in the capability of heteroduplexes of DNA molecules to bind to the matrix in the column, homozygous variation (unable to form heteroduplexes) can not be detected unless the samples are mixed in an equimolar proportion with control DNA samples, known to have no heterozygosity within the relevant amplicon, a highly impractical procedure.

Most of the coding variations found in our cohort are unique or at least rare. This may indicate that homozygous coding mutations are not likely to be significant contributors to the disease in our patients, with the exception of cases where consanguinity exist. Much more important to consider is the existence of compound heterozygous. No patient carrying two of the coding variations described above has been found. The search for a compound heterozygous subject in the future must include also non-coding variations.

In the only known instance of a mutation of *MYH6* causing CHD, an I820N variation seen in a large family with atrial septal defect (Ching et al. 2005),

the inheritance pattern was clearly autosomal dominant. It is therefore surprising that a stop codon and a mutation of a donor splicing site, found in affected subjects, were inherited for healthy parents, and in the case of the former, is present in an equally healthy grandfather. Incomplete penetrance could account for this situation, nevertheless, as the stop codon occurs much "earlier" in the protein than the missense "Mendelian" mutation, a dominant negative mechanism for the later could be proposed.

Dominant negative mutations often occur in parts of the protein capable to interact with other molecules. I820N is located in the myosin regulatory light chain binding site of MYH6.

In an alternative explanation, mutations affecting the function of the myosin monomer encoded by the defective allele, leaving intact the function of an interacting partner could behave as recessive mutations whereas mutations affecting interactions could produce disease in a dominant manner.

The limits between potentially Mendelian mutations and susceptibility alleles become blurred when a Mendelian or at least familial form of a mostly complex disease is being studied. As we do not have at the moment access to more samples from the extended family of the patient

with the nonsense mutation is not possible to obtain quantitative information about penetrance. The approach in this cause could be to track the mutation instead of the phenotype across the family.

3.6.1 Branch site mutation

Mutations of the "T" or "A" positions of the branch site consensus "ynytray" in certain genes have been shown to cause disease (Kralovicova et al. 2004). The one base pair deletion observed in 14 of the 380 CHD patients and in 4 of 192 control subjects occurs in the "T" position of one of the only two segments compatible with the consensus found in the whole intron upstream exon 14 of *MYH6*. The splicing alterations the *MYH6* transcript would depend on the existence and functionality of alternative potential branch sites in the same intron.

If there is not an alternative branch site in the same intron, the outcome of splicing depends on its size. In the case of short introns, there is a tendency towards complete intron retention, probably because the branch site of the intact intron is more likely to carry out the attack in its own splicing donor site (Kuivenhoven et al. 1996), For long introns the tendency favours exon skipping, when the branch site of the intact intron establishes the covalent bond with the phosphate group of the donor site

of the mutated intron (Burrows et al. 1998; Khan et al. 2004; Putnam et al. 1997).

Chapter 4 COPY NUMBER ANALYSIS OF *MYH6* IN PATIENTS WITH CONGENITAL HEART DISEASE

4.1 Introduction

Copy number variations (deletions, duplications) are a normal feature of the genome of healthy individuals (Conrad et al. 2006). It has been estimated that two given normal subjects typically show more that 11 copy number variation differences in sequences longer than 100kb (Sebat et al. 2004). Most of these variants seem to behave in a way similar to SNPs. They appeared once in evolution and show linkage disequilibrium with other surrounding Mendelian elements (Hinds et al. 2006).

In contrast, some recurring deletion and insertion events are responsible for numerous Mendelian disorders like Charcot-Marie-Tooth disease type 1A (Nelis et al. 1996) and DiGeorge syndrome (Tezenas Du Montcel et al. 1996). These recurring deletions are typically produced by non-allelic homologous recombination events between sequences longer than 10 kb,

50kb to 10Mb apart, located in tandem and with a similarity of 95% or higher (Inoue and Lupski 2002).

Different types of repetitive elements in the genome like *Alu* and LINES, are also known to mediate this kind of rearrangements. Alu elements seem to have originated as an accidental retrotrasposition of a 7SL transcript (Rowold and Herrera 2000) and are thought to have played a role in the divergence of humans and chimpanzees (Sen et al. 2006). It has been estimated that *Alu* recombination-mediated deletions and insertions account for 0.3% of human heritable disorders. These recombination events can occur intrachromosomally (producing insertions and deletions of variable length, often spanning entire genes or several exons in a gene) or interchromosomally causing chromosomal structural abnormalities (Deininger and Batzer 1999).

Non-allelic homologous recombination has been documented between the *MYH6* and *MYH7* genes. A 5'*MYH6*/3'*MYH7* hybrid gene flanked by intact copies of *MYH6* and *MYH7* was discovered in a family with hypertrophic cardiomyopathy (Tanigawa et al. 1990), although it was established that the hybrid was not related to the phenotype (Watkins et al. 1992). The recombination event from which the hybrid gene derived, necessarily produced in the same meiosis the counterpart rearrangement, a hybrid gene 5'*MYH7*/3'*MYH6* without normal flanking copies of *MYH6* and *MYH7* (see Figure 4.1), but such rearrangement has not been found.

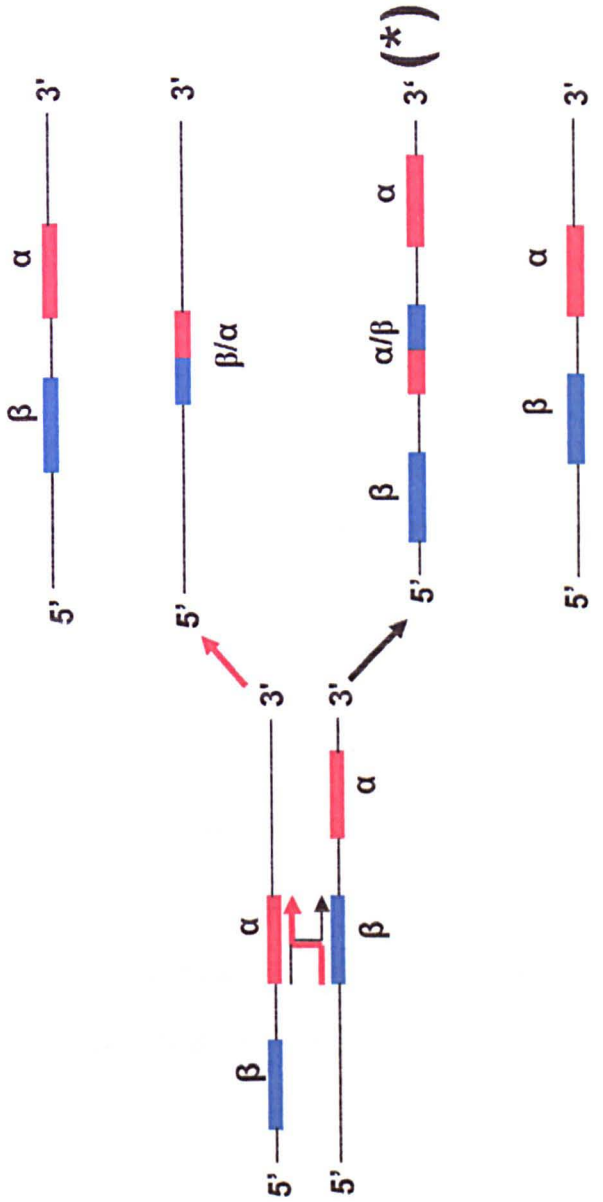


Figure 4.1: Non-allelic homologous recombination events between the α- (*MYH6*) and β- (*MYH7*) heavy cardiac myosin genes. Their high degree of similarity and close proximity makes them susceptible to these kinds of recombination events. Tanigawa, Jarcho et al (1990) reported a hybrid gene (*) 5'α/3'β flanked by two normal copies of α and β in a family with hypertrophic cardiomyopathy (black arrow). The other rearrangement produced by this recombination event, a 5'β/3'α hybrid without flanking α or β genes (red arrow) has not been found.

4.2 Multiplex Amplifiable Probe Hybridization

Multiplex Amplifiable Probe Hybridization (MAPH) is a method that allows the identification of copy number variations in a complex DNA sample and is based in the capability to amplify in a multiplex PCR reaction probes specifically bound to genomic DNA fixed to a nylon surface. The probes used are short (100 to 600 bp), and are flanked by the same pair of PCR primer binding sites. The length of every probe is different so they can be resolved and quantified by electrophoresis once amplified in the final multiplex PCR. The comparison between the amounts of different test and control probes recovered from the nylon filter, after stringent washes to eliminate not specifically bound probe, is a reflection of the amount of targets present in the DNA attached to the membrane (Armour et al. 2000).

The aim of the project described in this chapter was to use MAPH to answer the question of whether non-allelic homologous recombination between *MYH6* and *MYH7* or repetitive elements recombination-mediated deletions could account for the phenotype in some of the patients in our CHD cohort.

4.3 Materials and Methods

4.3.1 Design of the MAPH probes

The sequences of all 39 exons and their neighbouring parts of the corresponding introns were submitted to BLAST searches against the human genome in order to identify possible cross-hybridization targets, ie regions of the genome besides the exon whose sequences was used as a BLAST query with strong enough similarity to allow the binding of a probe covering the exon.

As MAPH enables us to easily detect 25% variations in dosage, that is the reduction expected in the case of a deletion of an two-copy element per haploid genome in a diploid organism, exons with no or one potential cross hybridization targets were used to make probes (exons 1, 2, 3, 6, 9, 10, 13, 16, 25, 36, 18, 27, 28, 30, 32, 34, 39 and a 408bp fragment of the promoter) whereas those with more than one potential cross hybridization target were screened using intronic probes flanking the exon (the remaining exons) (see Figure 4.2).

The sequence of the segments used for potential intronic probes were submitted to BLAST searches against the human genome in order to detect similarity with repetitive elements.

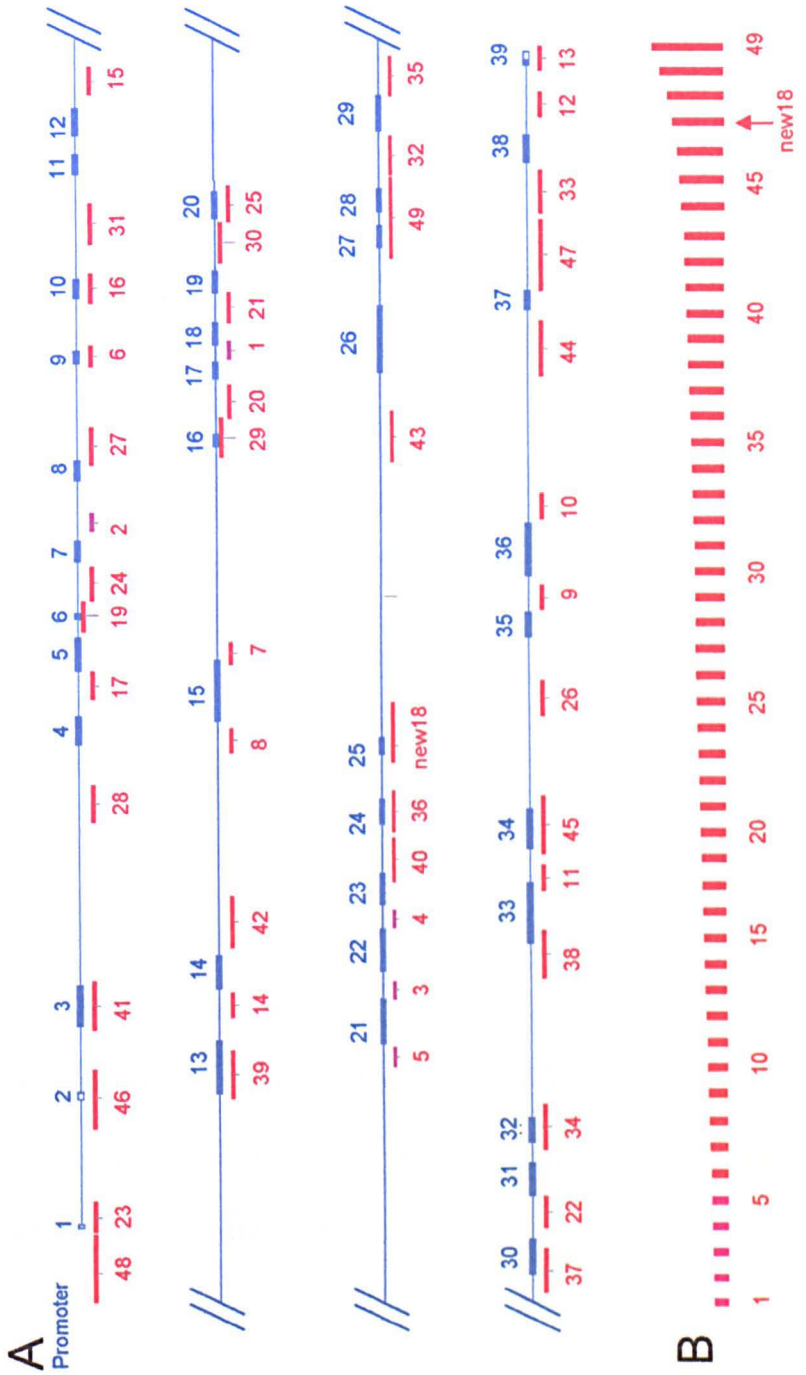


Figure 4.2: Scaled diagrams showing the genomic organization of *MYH6* and the localization and relative size of the 49 *MYH6* probes used for MAPH. A) Exons (blue boxes and numbers) are labelled according to their order 5' to 3'. Probes (purple and red boxes and red numbers) are labelled with respect to their size, except for "new18" probe whose size is between 46 and 47 probes (there is no 18 probe). B) Probes aligned by size.

In addition to 49 *MYH6* probes X (*ZIC3*) and Y (*SRY*) linked probes were designed in order to detect normal variations of gonosomes according to sex. Control probes in other autosomes (*MEF2C* in 5q14, *IRX3* in 16q12, *HEY2* in 6q21) were designed to be able to detect whole-gene deletions of *MYH6*. A non-human probe was designed as a control for non-specific binding and was derived from lambda-phage DNA, with a total of 55 probes (Table 4.1 and Box 4.1).

4.3.2 Preparation of plasmid stocks

Once all the probes were designed as amplicons (Figure 4.2), primary 30 cycles, 20µl PCR reactions were performed using genomic DNA from a control subject as template for each probe plus a zero DNA control. Secondary 20 cycles PCR reactions with a 0.5µl input from the primary PCR were carried out (Figure 4.3A).

Amplification was assessed using 2% agarose gel electrophoresis of 2µl of each secondary PCR reactions. DNA from the remaining 18µl was purified using ethanol precipitation and treated with the Klenow Fragment of DNA polymerase I to remove the template-independent 3' terminal Adenosine nucleotide added by *Taq* polymerase at the end of each polymerisation cycle (Figure 4.3B).

Table 4.1: Sequence of PCR primers used to produce amplicons for the MAPH probes. Sizes of the amplicons and full probes, as well as GC content of the amplicons are also shown. Capital letters indicate exonic sequence

Probe	Exon	Amplicon	Full probe	Primer 5'	Primer 3'	GC%
1	18(5')	89	148	aagaccttagtctggggaggac	GTTGTCCATCACCCctgtgtc	56%
2	7(3')	94	153	tgacccgagttacccctaacc	gtcttccatactgggctgacc	63%
3	21(3')	97	156	ggccacgtgattatctcttcag	aaaagagctggcactcctagac	55%
4	22(3')	100	159	cagctcccctctggcttcag	ccagcctggagacatctatgg	59%
5	21(5')	103	162	catgttcagtgtagtgatgttgg	ataaaagagagcttcccaatcatgg	48%
6	9	106	165	ctctgtcccattcgtccccag	agcctcoctgctggtactcac	55%
7	15(3')	109	168	CAACTGCCGATACTGgtaagc	cctgcctatggagtcatgtgc	59%
8	15(5')	113	172	aatattggcttcctgttttaggg	aggcttccacagtcccatacc	54%
9	36(5')	116	175	cccctaaagacagaaacaagg	cccaaggatctcctttctcc	55%
10	36(3')	120	179	gatgtggaagtttcttctctgg	aagtcaaactgactgcagagc	53%
11	33(3')	123	182	AAGCAAGTCAAGAGCCTCCAG	ggaggaatctggtgcctgtatc	60%
12	38(3')	127	186	cagttctgagggtcccatagc	gttagaggcacttgtggtttagg	52%
13	39	130	189	gccccctcacacctcttatt	aaggacagatgggctcagg	49%
14	14(5')	134	193	cagtcattgtccctgtcttcag	agcagtcagaaagtgggtgtg	62%
15	11-12(3')	137	196	agagactctgaggcttgtgg	cacctcagttctggtgagtgc	60%
16	10	141	200	cctcccctccacccacctcag	gggggtggcaggcaggttcac	55%
17	5-6(5')	144	203	actcccaagggacccaagt	aaggctgggcatgaggtt	63%
19	6	152	211	cagacacccacccagatcc	accctaggcatcagcgtgta	58%
20	17(5')	155	214	tgatgaggaaagaagctaggg	catcccttaggcccttaaacc	48%
21	19(5')	158	217	GGCAGAGgtgggtatgagg	CGATACctgaggagggaagtg	59%
22	31(5')	161	220	agaggccaaaggcaacctc	gacaaggctcactcttcagc	65%
23	1	166	225	taaaggaactggagctttgagg	ttaatagagcccttaggcctcc	59%
24	7(5')	169	228	gaggtttaaggctgggattgc	aggattcctactgttctccatacagg	56%
25	20	176	235	tctacccctgcctgcag	tcccttttcctcctgtctca	60%
26	35(5')	181	240	gacatcatccacttacctcatcc	atagcaaaaggtctgcagtattacc	41%
27	8(3')	185	244	AACTCCTCCCGCTTTgtgag	agtggaaagtcagacctttgg	51%
28	4(5')	189	248	tgagcaacagatacoctaaagc	gtacctttctcttgccagatcc	57%
29	16	197	256	gggctcctttatttccagc	ataggtggtgcagccagaag	55%
30	19(3')	202	261	atgttgagtggagcagagagg	gcagcctcagaggggagtatc	50%
31	11-12(5')	207	266	tgcagtcacacacctcactcc	agaagtcagcaaggaccaagg	57%
32	29(5')	211	270	aagagtcctctgcaagggaaag	ctgtgccccattctctagattc	49%
33	38(5')	215	274	cagcaaaacaggattctgaagg	gggattctcagcctctcaacc	56%
34	32	221	280	cataggctttgagctttctgg	gggagaagtgggagacagc	59%
35	29(3')	223	282	CAAgtgagctccagataccc	gtataacccgggccaaaagc	57%
36	24	230	289	tttcatgttttccacactttgc	gacacctccattagcccctc	51%
37	30(5')	233	292	cggtgtctcaaggagatatagg	CTCTACGTCCACCATCAAGTCC	59%
38	33(5')	241	300	cagcactttctcccaatagc	CCTTGATCTGGTTGAACTCTAGC	62%
39	13	245	304	tcacttatcctttcoctctcaac	GTGGCTGCTTGGTCTCCAG	58%
40	23(3')	250	309	CTGTCCAAGTCTAAGGTCAAGC	cttgaagtccagtgggattgg	47%
41	3	255	314	gccctcctgtctctgacccag	cggcgccatgccctactcacC	64%
42	14(3')	263	322	agacaaagtggtggctgagtcc	taggatctttccatgaatgacg	54%
	ZIC3	274	333	CCCTAGCTACTTGCTGTTTCC	GATAGGCTGCCGCATATAACG	66%
43	26(5')	280	339	tatggttgattggactttgtgg	taggccatcaataggaatttgg	45%
44	37(5')	287	346	gaacctatgtaagtccaggttgg	cctccatcattttactcttctcc	49%
45	34	293	352	ctgctcacacccactctcc	ctagatgtcctgggctctgc	63%
46	2	299	358	tctgactccctggtctgtcc	ctggagtatgctaagggttgg	59%
	SRY	305	364	CCATGCACAGAGAGAAATACCC	agcatctaggtaggtctttgtagcc	53%
	IRX3	321	380	ctcctgaatctcacctcttttgc	AACAAACCTCACAGCGAATGC	40%
	new18	335	394	acctgtcactccccacc	cccctctcaaacagttttgc	50%
47	37(3')	366	425	Ggtgagttcagagctttcttcc	tgctaatcagcaactcacatcc	48%
	MEF2C	374	433	ttcttcggaaacgtatttgacc	aagcaaggctctgtcaatgg	37%
	HEY2	391	450	cattttgctgtggtgatcttagg	tgaagtcatctacttacCTTTACCC	30%
	Lambda	400	459	gagagttaatttcgctcacttcg	gattcacaccgactcatttaagc	45%
48	Promoter	408	467	aaaggagaggctggggaac	cccaccctctgtctaaatttggagtc	58%
49	27&28	461	520	ccctctcttccttcctctgg	ctggcactgagatgaattgc	59%

Box 4.1: Sequences of the PCR amplicons used to produce the 55 MAPH probes. Capital letters indicates exonic sequence.

>1) MYH6 Exon 18 5', 91bp
aagaccttagtctggggaggacagctggcatccactttaccctaaggctgacccttttccccttccctcctgacacagGGGTGATG
GACAAC

>2) MYH6 Exon 7 3', 94bp
tgacccgagttacccctaaccctccctccctgtgacgtggtggggacagccacactgagctgggctcccgatggtcagcccagtat
ggaagac

>3) MYH6 Exon 21 3', 97bp
ggccacgtgattatctcttcagccctctcctcccctcccctagattatagcccatctcacaaccagggactgggagtctaggagtgcca
gctctttt

>4) MYH6 Exon 22 3', 100bp
cagctccctctggcttcagcccaggtctcctcaagactcccagactagagtgttgtcctggtccttggcatggaggtccccatagatgtc
tccaggctgg

>5) MYH6 Exon 21 5', 103bp
catgttcagtgtagtgatgttggaggtgggtgtgcacaagaaggaagtctacgtgcctacgaacttgcttagtagggcccatgattggg
aagctctcttttat

>6) MYH6 Exon 9, 106bp
ctctgtcccattcgtccccagGGGAAATTCATTAGGATCCACTTTGGGGCCACTGGAAAGCTGGCTTC
TGCAGACATAGAGACCTgtgagtaccagcagggaggct

>7) MYH6 Exon 15 3' 109bp
CAACTGCCGATACTGgtaagcaggcagcccctgcactgggccaggggacttctgaagacacaaagggccagggtcct
gctgcttcaaagcacatgactccataggcagg

>8) MYH6 Exon 15 5' 113bp
aatattggcttcctgttttagggtaagaggtaccagcacagcgccccttcagcaggccagcgctactggctccagattccttttcctgtca
gggtatgggactgtggaagcct

>9) MYH6 Exon 36 5' 116bp
cccctaaagacagaaacaaggccttgggtccaggccaggccactgtgctgtaacaccaagccaactctgcagttctgtggatttga
gggcctgatgggagaaaggagatccttggg

>10) MYH6 Exon 36 3', 120bp
gatgtggaagtttcttctctggccccactgccccgccctcacagggctcctctcacctcctccttgagatgctgttggtagatttaacgttctt
ctcacgctctgcagtcagtttgactt

>11) MYH6 Exon 33 3', 123bp
AAGCAAGTCAAGAGCCTCCAGAGCTTGCTGAAGgtacatgggggcgggaggtcccctcaggggactggcct
ccatgtggcctggagaagcagtggtgtctggatacaggcaccagattcctcc

>12) MYH6 Exon 38 3', 127bp
cagttctgagggtcccatagcttacataacctgagaatccactctcctgctcaaaacagccccccactgactggaacttctgcagaga
tccccagttccatccccctaaaccacaagtgcctctaac

>13) MYH6 Exon 39, 131bp
gccccctcacacctcttattcttttgcagCAAAAAATGCACGATGAGGAGTGACACTGCCTCGGGAACCT
CACTCTTGCCAACCTGTAATAAATATGAGTGCCAaactctgcctgagcccatctgtcctt

>14) MYH6 Exon 14 5', 134bp
cagtcattgtccctgtcttcaggggaagccctcctccactgcctgacatggaggggacagccataccctgctgggctcggcacagt
gcacgggcacagccccaatggccactcacacccactttctgactgct

Box 4.1(continued)

>15) MYH6 Exon 11-12 3', 137bp
agagactctgaggcttgtggggcagtggggtgtgttgtgggatggtgagggacagggacgtgggagacgtgggcagacagagagt
ccaccacaatcccaactcctcaaccccagctgcactcaccagaactgaggtg

>16) MYH6 Exon 10, 141bp
cctcccctccacccacctcagACCTGCTGGAGAAGTCCCGGGTGATCTTCCAGCTGAAAGCTGAGAG
AAACTACCACATCTTCTACCAGATTCTGTCCAACAAGAAGCCGGAGTTGCTGGgtgaacctgcctg
ccaccccc

>17) MYH6 Exon 5-6 5', 144bp
actcccaagggacccaagtccccttcccccctccaccccctagtcagctgcaggaggagtagagccagctggagtgaacagggaca
tgcctggctgccaccactgcctgtcccaggctctccccaccaacctcatgcccagcctt

>19) MYH6 Exon 6, 152bp
cagacacccacccagatcctcagccctgaccccattgcttctcctctttttttcttccagATCGGGAGAACCAGTCCATCCTC
ATCACgtgagcgagtgccatcctcccacagaagggactgggctgggggcatacacgctgatgcctagggt

>20) MYH6 Exon 17 5', 155bp
tgatgaggaaagaagctagggctacgtagtcgtttttagtgtgtagattccacaaaagcctgaactcacgtcacaaattatgtggcagc
ccctgtccctttgatgattctggtggagagtgtctggggaccagggtttaagggcctaagggatg

>21) MYH6 Exon 19 5', 158bp
GGCAGAGgtgggtatgagggtgccccagagctcatagaacaggggggagccaggctgccctgatgggaatgggatctgcaggt
gaccctggaattctgtgggcagagcagatcactgcagagcatgggtgactctggacacttccctcctcagGTATCG

>22) MYH6 Exon 31 5', 161bp
agaggccaaaggcaacctccttggaggtggaggaggagggctaagcccaggctcgggaccagggacagatcttggacatgcgg
ctgaggctgggggctggggcactgggaagcaggagggctggggagctaaggctgggggctgaagagtgagccttgtc

>23) MYH6 Exon 1, 165bp
taaaggaactggagctttgaggacagATAGAGAGACTCTGCGGCCCAGgtaagaggaggtttggggtgggatgccct
gcagcccgtccacagagcccccaccgtgagggacctccttcaccaggagtggggtgcaggtcagttggaggcctaagggctctatt
aa

>24) MYH6 Exon 7 5' 170bp
gaggtttaaggctgggattgcagggagcatggggcactgagctctcattagaggggtgccagagcaatgaaccatgtcaggcaaat
cctgctgggtgctgggctctgatggccaggctgggaaggggagcatgtgatgctgagccctgtatggagaacagtaggaatcct

>25) MYH6 Exon 20, 176bp
tctacccctgcctgcagGTGTTCTTCAAGGCAGGGCTGCTTGGGCTGCTGGAGGAGATGCGGGAT
GAGAGGCTGAGCCGCATCATCACGCGCATGCAGGCCCAAGCCCGGGGCCAGCTCATGCG
CATTGAGTTCAAGAAGATAGTGGAACGCAGgtgagacaggaggaaaaggga

>26) MYH6 Exon 35 5', 181bp
gacatcatccacttacctcatcccctcttccttaccccactacattcttggtatagctataaacatctctggaatattgtcctggtaaaaagtt
gtttccatcttccttggaatattgtcatgctcctacgtaaacatgttgctaaagagctccagggtaatactgcagacctttgctat

>27) MYH6 Exon 8 3', 185bp
AACTCCTCCCGCTTTgtgagtgcctttgaccactcccagtggcctcatccagccttgacaagaaaaaggggtgctgttttgc
cacacccagttgattgtactgtatctggctttgggatatcaacatgcacacgttggtgggaagagagcataggctttggagccaaaggt
ctgactttccact

>28) MYH6 Exon 4 5', 190bp
tgagcaacagataccctaaagcgctgcccgcgggagacagcctcggggtcagcataaggtgtgcacagatctgagagctgccaat
ctccaggtctgccccaagaccttggaacatagggactgaagagtgatggtcatgggcacagggtgtccccaggatggtctggggat
ctggcaagagaaaggtac

Box 4.1(continued)

>29) MYH6 Exon 16, 197bp
gggctcctttattttccagctcctgttgattatttctcctctcgctgtttagGGGACAGTGGTAAAAGCAAAGGAGGCAAG
AAAAAGGGCTCATCCTTCCAGACGGTGTCGGCTCTCCACCGGgtaagaagggcccaggggtgccag
gacacctggtggaatggcccagcccagagacttctggctgcaccacctat

>30) MYH6 Exon 19 3', 202bp
atgttgagtggagcagagaggagtttaggaggcagaagcctaattctggcttccttatcaaccttatcaagggctgaaacccaggcttc
attccggtcttgtttgtcaaatttttactcttacttctagaaggcatggggtgatgggtcacctgggagctcatccagggtcttccaccctgg
atactcccctctgaggctgc

>31) MYH6 Exon 11-12 5', 207bp
tgcagtcacacacctcactccttctccaccctctctccactcctccctcggcttgtctttgccttctcctctttatttgcctccccatctcccatgt
gtcactcccctgccaccctccctcctccttgttactggagcccctgccagggcccctcttctaccctttcacttcctctgggttcactccttggt
ccttgctgacttct

>32) MYH6 Exon 29 5', 211bp
aagagtcctctgcaagggaaagaccctccagtctaggttctgccctgcagctaagcgtcatttaatgcctcttttcttattcgtaaggggat
ggggtgagcagactgggaaactcctcaaacagtgaggtgccacatcagcccacatggtgaataaggctgggcttggttgaagtact
acataagaagagaatctagagaatggggcacag

>33) MYH6 Exon 38 5', 216bp
cagcaaaacaggattctgaaggggcccagatcgggcagcatgggatttgtctggggcagtggatggccgtgaaggactctgagtg
ctggacatgtttgagaagagtgcaaggcagttgcaggataccttgggaaggctgttgcaggaatatgcatgaggtatgggtgcctca
gggacagggagctggaacctcaggttgagaggctgagaatccc

>34) MYH6 Exon 32, 221bp
cataggctttgagctttctggccctctggtccccagAGGAAATCTCGGACCTTACTGAGCAGCTAGGAGAAGG
AGGAAAGAATGTGCATGAGCTGGAGAAGGTCCGCAAACAGCTGGAGGTGGAGAAGCTGG
AGCTGCAGTCAGCCCTGGAGGAGGCAGAGgtgagggccgagaactccctgcaccccatccctgttctgccgctg
tctcccacttctccc

>35) MYH6 Exon 29 3', 225bp
CAAgtgagctccagatacccccttaacctgactctcagagaggaaggggcgagaggacctggggtggggacaggcaaagtggt
catgagacggaagtggaagagacaggaggaactcggagggcaacagaagtgcttggaagaaagcctgaactctttgctctgtga
actctggctggccctgacccacttcctgtgacgggcgagcttttggcccgggttatac

>36) MYH6 Exon 24, 230bp
tttcatgttttccacactttgcttatttttcttccctccaacagCTGGAGGGATCCCTAGAGCAAGAGAAGAAGGTGC
GCATGGACCTGGAGCGAGCAAAGCGGAAACTGGAGGGCGACCTGAAGCTGACCCAGGAG
AGCATCATGGACCTGGAAAATGATAAACTGCAGCTGGAAGAAAAGCTTAAGAAgtaggagactgt
ggtggccaggaggggctaatggaggtgtc

>37) MYH6 Exon 30 5', 234bp
cggtgtctcaaggagatatagggaggggtggaaggaggtccacccaaggctccagtgttgcccagtagagtcacacacacacc
ctccaccctcacctgggcagAAAGAAGCTGGCCCAGCGGCTGCAGGATGCCGAGGAGGCCGTGGA
GGCTGTTAATGCCAAGTGCTCCTCACTGGAGAAGACCAAGCACCGGCTACAGAATGAGATA
GAGGACTTGATGGTGGACGTAGAG

>38) MYH6 Exon 33 5', 241bp
cagcactttctccccaatagcaccgtggtgggcatggggtgcagatcccacctccccatctcttctcagctcttcttctctgggcgatagtc
ctggctgacaccgtgtatcttctcatcctccctctcaaccctgccctgtgccctgtctgcccgccctcgccccacccctcccagGCCT
CCCTGGAGCACGAGGAGGGCAAGATCCTCCGGGCCCAGCTAGAGTTCAACCAGATCAAG
G

>39) MYH6 Exon 13, 245bp
tcacttatcctttccctctcaaccagATGCTGACAAGTCGGCCTACCTCATGGGGCTGAACTCAGCTGAC
CTGCTCAAGGGGCTGTGCCACCCTCGGGTGAAAGTGGGCAACGAGTATGTCACCAAGGG
GCAGAGCGTGCAGCAGGTGTACTACTCCATCGGGGCTCTGGCCAAGGCAGTGTATGAGAA
GATGTTCAACTGGATGGTGACGCGCATCAACGCCACCCTGGAGACCAAGCAGCCAC

Box 4.1(continued)

>40) MYH6 Exon 23 3', 250bp
CTGTCCAAGTCTAAGGTCAAGCTGGAGCAGCAGGTGGATGATgtgagtagtaagaaccatgctcct
gctctcagagcaagattttgcaggcaacaccaatggcccagaaagtcctgatccctagaattaacttctatggcccctgaagctttt
ttgctctctgtagttcctcactacagtaggtctctgaatcctttgtgcttgcaggatttctctgttggtttgacttccaatcccactggacttca
ag

>41) MYH6 Exon 3, 255bp
gccctcctgtctctgacccagGGGAAGCACCAAGATGACCGATGCCCAGATGGCTGACTTTGGGG
CAGCGGCCCAGTACCTCCGCAAGTCAGAGAAGGAGCGTCTAGAGGCCCAGACCCGGC
CCTTTGACATTCGCACTGAGTGCTTCGTGCCCGATGACAAGGAAGAGTTTGTCAAAGCC
AAGATTTTGTCCCGGGAGGGAGGCAAGGTCATTGCTGAAACCGAGAATGGGAAGgtgagt
agggcatggcgccg

>42) MYH6 Exon 14 3', 263bp
agacaaagtggtggctgagtcccttctacacccaagaaactagagtcccaagaatcccaggccttctccaggcccagcttctcc
ccactgtgaagtcatgggcatgaacaggatgatccccccactcttcctttcccaggaccttgcactttatgcccctttgtggtggtccc
ctcagtgtcttaagagtgagatgtagtgaaggagaggcccctggcccctctgaccgcccatgagaagcgtcattcatggaaagat
ccta

>ZIC3 274bp
CCCTAGCTACTTGCTGTTTCCCGGGCTGCATGAGCAGGGCGCTGGGCACCCGTCGCCC
ACAGGGCACGTGGACAACAACCAGGTCCACCTGGGGCTGCGTGGGGAGCTGTTCGGC
CGTGCTGACCCATACCGCCCAGTGGCCAGCCCGCGCACGGACCCCTACGCGGCCGGC
GCTCAGTTTCCTAACTACAGCCCCATGAACATGAACATGGGAGTGAACGTGGCGGCCCA
CCACGGGCC CGGCGCCTTCTTCCGTTATATGCGGCAGCCTATC

>43) MYH6 Exon 26 5', 280bp
tatggttgattggactttgtggttaacttggagaattgcaaaggtatctgattgtttcgaggcatgttgtcacaaatatttgtaaaatac
agcactcatttttcccgtcttatgaatagcgcaacagagcctagtgaatctggggactctgaacttcttgatctcacaggataccagg
atccccctcaaccacaggttctcaggatttggggctgcagatgctcacactgggtctgagatgcccttgggagcttcagccaaatt
cctattgatggccta

>44) MYH6 Exon 37 5', 287bp
gaacctatgtaagtccaggttggagctttatccaccatactggagctggaacagacctggtgcttttatattaccacattagggaatt
ccattaggttctgagcccctcccctacttctagctttatgacttcagccttcattgctctgtggatccctgactgacaaccttgcattgcc
cctttgacctacgatagagtcagagaatcttccccaccacctctttgacctggatcattgcagggaggggcagcaaaggcaagg
ggagaagagtaaaatgatggagg

>45) MYH6 Exon 34, 293bp
ctgctcacacccactctcctgatgctcagGACACCCAGATCCAGCTGGACGATGCGGTCCGTGCCAAC
GACGACCTGAAGGAGAACATCGCCATCGTGGAGCGGCGCAACAACCTGCTGCAGGCTG
AGCTGGAGGAGCTGCGTGCCGTGGTGGAGCAGACAGAGCGGTCCCGGAAGCTGGCGG
AGCAGGAGCTGATTGAGACCAGCGAGCGGGTGCAGCTGCTGCATTCCCAGgtgaggggtc
aggagccaccttgtggaaacctactgagtgcagagcccaggacatctag

>46) MYH6 Exon 2, 299bp
tctgactccctggtctgtcctgcctgtctgcgctctggggctgcctccatccccgggtggcctgcctctgttgttcttcactctcctcatctg
ttcttctctctgcccggctctacctctgttgttccttgctccacccaCGGTCCAGATTCTTCAGGATTCTCCGTGAA
GGGATAACCAGgtgagaactgcccccatttttctctgcagagactggggcatgcttctcctgggagccggattgctggacca
ggggtctgctgtcccaagcactcagcgccaacccttagcatactccag

>SRY, 305bp
CCATGCACAGAGAGAAATACCCGAATTATAAGTATCGACCTCGTCGGAAGGCGAAGATG
CTGCCGAAGAATTGCAGTTTGCTTCCCGCAGATCCCGCTTCGGTACTCTGCAGCGAAGT
GCAACTGGACAACAGGTTGTACAGGGATGACTGTACGAAAGCCACACACTCAAGAATGG
AGCACCAGCTAGGCCACTTACCGCCCATCAACGCAGCCAGCTCACCGCAGCAACGGGA
CCGCTACAGCCACTGGACAAAGCTGTAGgacaatcgggtaacattggctacaaagacctacctagatgct

Box 4.1(continued)

>IRX3, 321bp
ctcctgaatctcacctcttttgctactactgtctctctcttccagGCCCCAGAACCATCTGGACGCCGCCCTGGTCT
TATCGGCTCTCTCCTCATCCTAGTTCTTTAAAAAAAACAAAAAAACAAAAAAAACTTTTTTTA
ATCGTTGTAATAATTGTATAAAAAAAATCGCTCTGTATAGTTACAACTTGTAAGCATGTCCG
TGTATAAATACCTAAAAGCAAAACTAAACAAAGAAAGTAAGAAAAAGAAATAAAACCAGTC
CTCCTCAGCCCTCCCCAAGTCGCTTCTGTGGCACCCCGCATTCGCTGTGAGGTTTGTT

>new18) MYH6 Exon 25, 335bp
acctgtcactccccacccccaccttctcctgcagGAAGGAGTTTGACATTAATCAGCAGAACAGTAAGAT
TGAGGATGAGCAGGTGCTGGCCCTTCAACTACAGAAGAAACTGAAGGAAAACCAGgtgacttt
ttttcccagtgcatgaaagtgggagctcaatagccctgaggtaactgaggctgcagcagctgcttagggttctacacaatatctgga
actccaggcagcctcatagacccaaccatccctgacttacaggcgctcaggaacactagccttcccccatagagcaagaatacat
tacgttagcaaaactgtttgagaagggg

>47) MYH6 Exon 37 3', 366bp
Ggtgagttcagagctttcttccctttctcatcaacacacctactatttgtgagaaccaatgaatatctcctacagaggggcctggacaa
agagtttgctataaactttaactctcaaacatttgtttgacacatctggtatgctcagagctgtcaggtgttctgaattaacaaaggcacc
acctacaagctgcttacaattcagataccataacccaacagaaggcagtgtagatgctagtgcaggacgtggggcagccgaagc
tgaaggcaggaaggtggcggtggaattgggcctagaaggggacccagctagccacagggcagtggggaagacatctgggtg
gatgtgagttgctgattagca

>MEF2C, 374bp
ttcttcggaaacgtatttgacctagatttcttcctaaatgctttattgtttttgttaacagGGAACGGGTATGGCAATCCCCG
AAACTCACCAGGTCTGCTGGTCTCACCTGGTAACTTGAACAAGAATATGCAAGCAAAATCT
CCTCCCCCAATGAATTTAGGAATGAATAACCGTAAACCAGATCTCCGAGTTCTTATTCCAC
CAGGCAGCAAGAATACGATGCCATCAGTGgtaatacaaaactacattttaaataaatattgataatgttttgtatat
gttttgctgtttttatatttgtctaacttgatgaattacttaacattctcccagaaaattaacataacagtgccattgacagagccttgct

>HEY2, 391bp
cattttgctgtggtgatcttaggagaataattattcttatgatttaaaaaagtcattctagtttcttttagtctttaaaagttaaaatggcaaat
acttgtcattaggtatgacattaattagtgctttcatgtaaagtttggaaagaattcgaagttttctgtgaaaaaataaaagttgggattgt
ctagtgagagggaaattgaagccataatgctaaattcagtgtataatgtttcagattagtttctgtaggacataggattatttatttatatt
tatttatgaagGGATCTGCAAAGTTAGAAAAAGCTGAAATATTGCAAATGACAGTGGATCATTTG
AAGATGCTTCAGGCAACAGGGGGTAAAGgtaagtagatgacttca

>Lambda Phage, 400bp
gagagttaatttcgctcacttcgaacctctctgtttactgataagttccagatcctcctggcaacttgcacaagtccgacaaccctgaac
gaccaggcgtcttcgttcatctatcggatcgccacactcacaacaatgagtggcagatatagcctggtggttcaggcggcgcattttt
attgctgtgttgcgctgtaattcttctatttctgatgctgaatcaatgatgtctgccatctttcattaatcccgaactgttggttaatacgcttg
agggtgaatgcgaataataaaaaaggagcctgtagctccctgatgattttgcttttcatgttcatcgttcctttaaagacgccgtttaacat
gccgattgccaggcttaaatgagtcggtgtgaatc

>48) MHY6 Promoter, 408bp
gaaggagaggctggggaacaggggaggaaagcccatggttgggaggcggaggacaggcatttggcctgcaggagaaggtg
accctcacccatgttttcagttcaccctcgggttaaaaataactgaggtaagggccatggcagggtgggagaggcggtgtgagaa
ggtcctgtcttcccactatctgctcatcagccctttgaaggggaggaatgtgcccaaggactaaaaaaaggccgtggagccagag
aggctggggcagcagaccttttcttgggcaaatcaggggccctgctgtcctcctgtcacctccagagccaaaggatcaaaggag
gaggagccaggaggggagagaggtgggagggagggtccctccggaaggactccaaatttagacagagggtggg

>49) MYH6 Exon 27-28, 461bp
ccctctcttccttcctctggtcgactcagcccctcccacactcaccctcctgtcttgcttcctgaagGCAAACCTGGAGAAAG
TGTCTCGGACGCTGGAGGACCAGGCCAATGAGTACCGCGTGAAGCTAGAAGAGGCCCAA
CGCTCCCTCAATGATTTCACCACCCAGCGAGCCAAGCTGCAGACCGAGAATGgtgggtgcccc
taaccaaccccctgcctagggcaggacatgacttgtgaaatggcccacaagcccctcatttcacctccagGAGAGTTGGCC
CGGCAGCTAGAGGAAAAGGAGGCGCTAATCTCGCAGCTGACCCGGGGGAAGCTCTCTTA
TACCCAGCAAATGGAGGACCTCAAAAGGCAGCTGGAGGAGGAGGGCAAGgtgaggcccagtg
gggagggtgggcaggcttgatggcagccctggggcaattcatctcagtgccag

The Klenow digestion reactions were loaded in a 1.5% agarose gel and DNA was purified from agarose blocks cut from the gel containing the fragments corresponding to each one of the 55 probes and re-dissolved in 30µl of water using the QIAGEN gel purification kit, according to the manufacturer's directions.

5.5µl aliquots of the purified DNA solutions were used in 10µl ligation reactions that included 100ng of pZero2 vector linearized with EcoR V (Invitrogen) and 2U of T4 DNA Ligase (New England Biolabs). The ligation reactions were incubated overnight at 16°C.

2µl of each ligation reaction were used in $CaCl_2$ competent TOP10 E. coli transformation reactions. Two Petri dishes containing 50µg/ml Kanamycin LB-agar for each one of the 55 transformation reactions plus a negative (no insert added) and positive (a known pZero2 recombinant plasmid) were inoculated and incubated overnight at 37°C in order to obtain isolated bacterial colonies.

Material from 2-4 colonies per transformation reaction was used to inoculate 5ml 50µg/ml Kanamycin LB-Broth overnight cultures (37°C in orbital incubator at 200rpm). Plasmid DNA was prepared for each

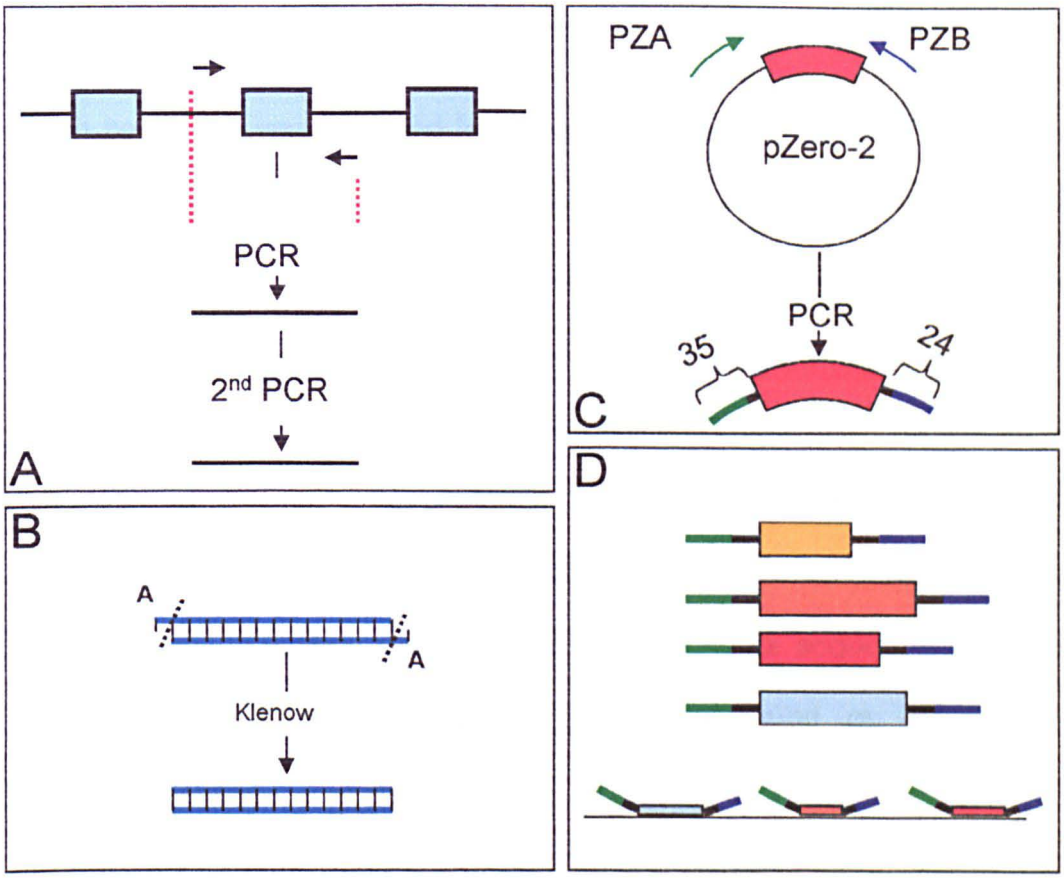


Figure 4.3: MAPH probe preparation. A) The desired genomic segment is amplified by PCR. Large amounts of DNA are produced by means of a secondary PCR reaction. B) The Adenosine residues incorporated by *Taq* polymerase at the 3' of each molecule are removed by Klenow fragment of DNA polymerase I digestion in absence of dNTPs. C) The "blunt" PCR product is the cloned into the EcoRV site of the pZero-2 vector. The plasmid is used as template of a PCR reaction with vector specific primers PZA and PZB in order to produce the probes. D) Each probe has a variable segment (boxes) that binds the genomic DNA fixed to the filter, and two peripheral constant segments derived from the vector used as binding sites for the primers used for multiplex amplification.

overnight culture using the Qiagen Plasmid Miniprep Kit and sent to sequence with M13 primer to our central facility.

Each sequence was received from the central facility as "seq" (sequence in 4-letter code) and "ab1" (electropherogram) via e-mail. Those files were used as queries to BLAST comparisons against the sequence of the genomic segment spanned by the probe. Those plasmids showing a sequence corresponding to the intended segment were selected as templates for probe PCR production. For a fraction of transformations, no adequate plasmid was identified by sequence. In those cases additional colonies were screened and plasmid selected if the sequence data was positive. Several probes required a second round of ligation and transformation.

4.3.3 Probe mix preparation

Once a set of appropriate plasmids was established for the 55 probes, the DNA concentration of the corresponding mini-preparations was measured using the Nanodrop Spectrophotometer (Nanodrop Technologies). All the plasmids were diluted to final concentration of 4ng/µl. 1µl of this dilution was used in a 20 cycles, 20µl PCR reaction with pZero-2 vector-specific primers PZA and PZB (Armour et al. 2000) in order to produce the probes (Figure 4.3 C).

Amplification was assessed running 5µl of the PCR reaction in a 2% agarose gel. 10µl of each one of the 55 PCR reactions were pooled in a single tube, and DNA purified from the mixture using the QIAGEN PCR purification kit.

The concentration of the purified mixture of DNA probes was measured and the solution diluted to a final concentration of 110ng/µl (2ng/µl for each probe). The probe mix was divided in 20µl aliquots and stored at -20°C.

4.3.4 Preparation of the filters

The samples used for the MAPH analysis consisted in genomic DNA solutions at different concentrations from 380 patients affected with different variants of congenital heart disease. For samples with a DNA concentration <100 ng/µl the solutions were concentrated using a SpeedVac centrifuge. 1µl of 1M NaOH was added to every solution containing 1µg of patient genomic DNA.

For each of the 380 samples a 4mm square filter was cut from a hybridization nylon membrane (Osmonics) and numbered with a sharp pencil. To the unlabelled surface of the filter, 1µg of NaOH denaturated genomic DNA was applied 1µl of solution at a time, letting the filter dry between applications.

Both sides of the filters were irradiated with 50mJ of UV light using a GS Gene Linker (Biorad) in order to bind the DNA to the nylon surface. Groups of 10-15 filters were introduced to an Eppendorf tube containing 1 ml of pre-hybridization solution and incubated 2 hours at 65°C. The solution was then replaced with a mixture previously boiled for 2 minutes of 200µl of pre-hybridization and 2µl of human Cot-1 DNA solution. The reaction incubated at 65°C for 30 minutes.

A mixture of 1µl of 110 ng/µl probe mix, 1µl of Cot-1 DNA (1mg/ml, Invitrogen), 1µl Herring sperm DNA (10mg/ml, Invitrogen), 2µl PhiX174/HaeIII (250µg/ml, Invitrogen), 1µl blocker mix (20mM PZAX and PZBX primers), and 2µl 1M NaOH was incubated at 37°C for 1 minute to denature the DNA molecules, the tube with the mixture was then placed on ice and 3µl of 1M NaH_2PO_4 was added and mixed. The solution was transferred to the tube containing the filters and the hybridization reaction was incubated at 65°C overnight.

4.3.5 Post-hybridization washes

The hybridization solution was removed after the incubation and replaced with 1ml of solution 1 (1xSSC, 1% SDS, pre-warmed and equilibrated overnight at 65°C). The tube was inverted several times. The solution was replaced again repeating the process. The solution was discarded and the

filters from 3-4 Eppendorf tubes were transferred to a 50ml Falcon tube (no more than 50 in total). The filters were washed 4 times adding and replacing solution 1 and inverting the tube repeatedly. After that 4 successive washes each one with fresh solution 1 were performed in a rotating hybridization oven at 65°C for 3 minutes. The solution from the final 3 minute wash was discarded and replaced with solution 2 (0.1xSSC, 0.1% SDS, pre-warmed and equilibrated overnight at 65°C) and the tube was inverted repeatedly to mix (quick wash) repeating the process once. The solution 2 was replaced again and the tube was placed in a rotating hybridization oven at 65°C for 5 minutes (long wash). Series of 2 quick washes, 1 long wash, 2 quick washed and finally 2 long washes were carried out after that (Figure 4.4A).

4.3.6 Recovery and PCR amplification of the specifically bound probes

Most of the solution 2 was discarded after the last long wash, pouring the rest along with the filters in to a Petri dish. From there they were transferred to 200µl PCR tubes containing 50µl of 1x PCR buffer. The tubes were then heated to 95°C for 5 minutes in a thermal cycler to release the probes from the genomic DNA fixed to the filter in to the 1x PCR buffer solution.

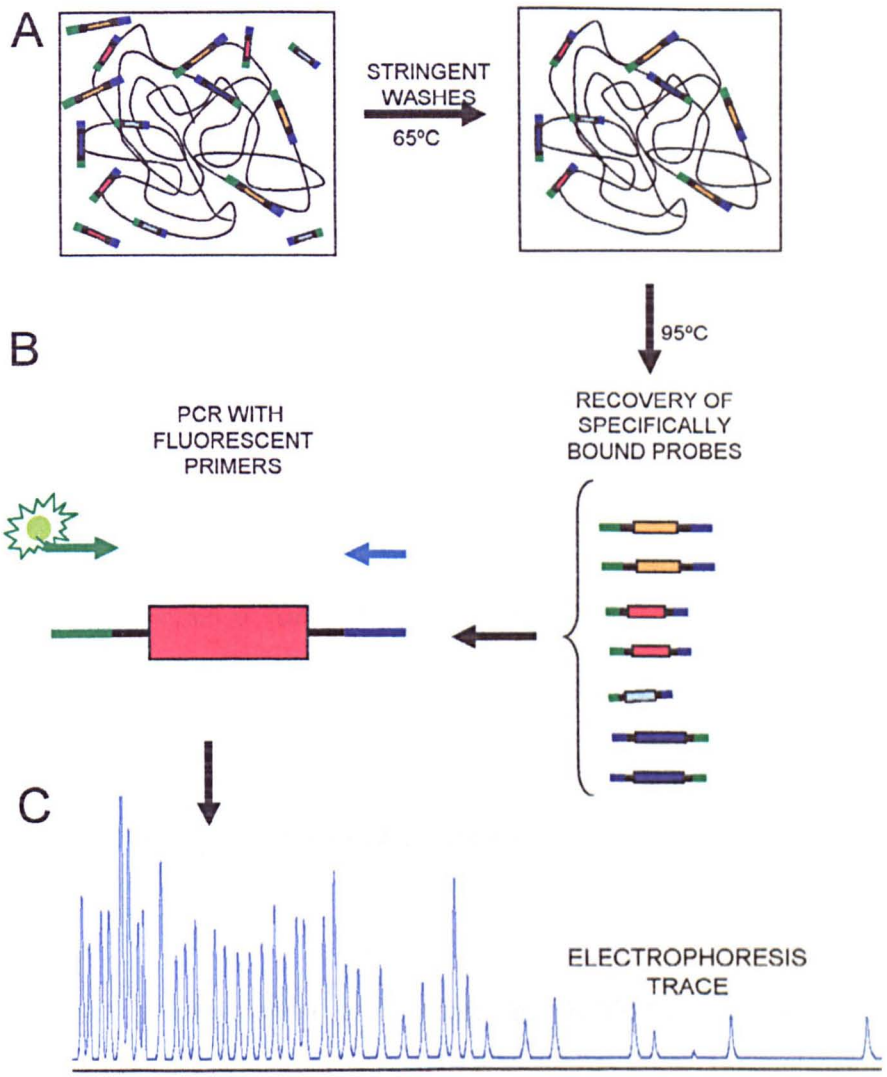


Figure 4.4: A) After hybridization the filters are washed several times with stringent solutions to remove probes non-specifically bound to the filter. B) The probes are recovered by boiling the filter in 1X PCR buffer. The solution is used in a PCR reaction with fluorescent PZA and PZB. C) Electrophoresis and fluorescent detection was carried out in an ABI3100 Genetic Analyzer.

1µl of the probe solution was used in a 20µl PCR reaction with 0.8µl of 10µM FAM labelled PZA primer and 2µl of 10µM PZB, 95°C 30 seconds, 60°C 1 minute, 70°C 1 minute for 25 cycles and an extension of 20 minutes at 72°C (Figure 4.4B).

4.3.7 Electrophoresis of fluorescent labelled PCR products

1.5µl of each one of the PCR reactions was added to 10µl of the mixture of 170µl of HiDi deionized formamide and 2µl of ROX 500 size standard (both from Applied Biosystems) in 96-well plates that were placed in a thermal cycler at 95°C for 3 minutes to denature the DNA duplexes and then on ice until ready for the electrophoresis.

Electrophoresis was carried out using an ABI3100 Genetic Analyzer (Applied Biosystems) using the run module Genescan POP4_30 and the analysis module GS500Analysis.gsp (Figure 4.4C).

4.3.8 Production of a size standard file for *MYH6* MAPH

MAPH traces from 40 control individuals were obtained to assess the reliability of the *MYH6* probe set and washing conditions. One of them was selected to produce a .szs standard size file that allowed automated recognition and analysis of the electrophoretic patterns of the samples by

the GeneScan software (Applied Biosystems). A custom .gta file for the analysis of the .fsa files corresponding to a single sample was produced using the Genotyper software (Applied Biosystems).

Through Genotyper, sample data (size and area of the relevant peaks in the electrophoresis traces) were extracted from the .fsa files originated by the ABI3100 and processed by GeneScan. A MS-DOS tab delimited table file was exported from Genotyper and imported in Microsoft Excel format .xls. Each row in this initial Excel table corresponded to an individual peak of an individual sample. The first column (labelled "Category") indicated the relevant probe, the second indicated the name of the relevant .fsa files and the third "Label", indicated the area of the relevant peak in arbitrary units.

The macro Rearranger (John Armour) was run through the Excel table in order to construct, in a separate "sheet" of the file, a new table with columns containing the areas of the peaks from each sample and rows showing the areas corresponding to each probe in different samples. A sex identifier was manually typed at the top of each column according to the gender of the individual tested and probes codes located at the left of each probe identifier to distinguish control, test, X and Y linked probes so the Maphematica macro could process them accordingly (see Figure 4.5).

Figure 4.5: Representative example (22 samples) of data in an Excel table constructed by the Rearranger macro.The first row was typed manually with identifiers according to the sex of the studied subjects (F, female; M, male). The second row indicates the sample number in the particular batch. In the first column a probe code was located according the type of the probe in each row (1, control probe; 2, test probe; 3, X-linked, control probe; 4, Y-linked probe). The second column contains the names of the probes. The numbers in the cells show the area of the peaks in arbitrary units as they were called by GeneScan.

Figure 4.6: Excel table constructed by the Maphematica macro. The areas of the peak in the first table (Figure 4.5) are replaced by the number resulting from the division of the area of each peak by the sum of the areas of the 4 nearest peaks in the same sample (relative ratios). The last column is constructed by the means of individual values of each row.

Figure 4.7: Second Excel table constructed by the Maphematica macro. The relative ratios of the first Maphematica table (Figure 4.6) are replaced by the result from a division of these values by the mean of the relevant row in the same table. The last column indicate the standard deviation of all the values in each row.

Probe Code	Number	F 2	F 3	F 6	M 7	F 11	M 12	F 13	F 16	M 17	M 18	F 19	M 20	M 21	M 23	M 24	F 25	M 27	M 28	M 29	F 30	F 31	F 32
1	P01	8957	10752	8242	20815	11748	9074	8127	10627	11285	8194	19136	10006	9353	9363	8148	10771	7690	10392	7652	14446	20951	19397
1	P02	8074	7269	4974	11872	7528	5341	5145	7181	7926	4997	12277	7291	5766	5773	5375	7033	5162	6859	5304	9413	10923	9937
1	P03	7350	9301	6814	15567	9361	6960	6442	8781	9722	6716	15846	8825	7803	8061	6633	9295	6495	8869	6528	11977	15197	13651
1	P04	7853	9746	7976	17817	10990	7626	7192	10005	10821	8038	17116	10005	8908	8955	7427	10722	7258	10335	7930	12451	16373	16045
1	P05	7579	9448	6982	17333	9979	7380	6642	9091	9712	7818	16447	8888	8246	8639	6693	9140	6810	9198	7157	12818	16891	16242
1	P06	6507	8532	6865	14055	8630	7257	6692	8600	8976	7552	14704	8500	8610	7519	6086	8958	6671	8464	6719	11460	14262	13323
1	P07	8023	10448	8129	19050	11290	8072	7367	10520	11127	7978	18102	9809	9606	9030	7579	9942	7873	10098	7424	13621	18390	17406
1	P08	6775	8937	6538	14796	9733	6906	6138	8453	9775	7597	16229	9064	8649	7737	6041	9101	7033	8408	6848	11949	15604	14742
1	P09	8041	9757	7998	14753	10038	8034	7325	9206	10616	8554	14780	9850	9661	8667	7447	10821	7658	10251	7870	12486	15249	14167
1	P10	7577	9570	7605	14577	9289	7485	7202	9142	10111	7705	15222	9025	8878	8347	6923	9763	6488	9681	6963	12111	15043	13111
1	P11	8109	10605	8774	17920	10934	8290	7974	10442	11881	7968	18419	10879	9306	9139	7767	10961	7134	9596	8012	13913	18666	15921
1	P12	6101	7956	5958	12740	8296	6410	5927	8087	8420	6478	13511	7900	7559	7022	5878	8184	5689	8138	6362	10318	12917	11152
1	P13	6189	7867	6221	11906	7880	6627	5930	7823	8164	5971	12110	7462	7151	6891	5506	8024	5680	7449	5963	9919	12282	10925
1	P14	16176	20343	15450	35965	22283	16409	14527	19683	21329	16771	35590	19979	19007	17739	14799	20445	15609	20168	15394	27306	36545	33151
1	P15	10804	13366	11002	22389	13668	11101	9778	13074	14230	10804	22704	12706	12736	11001	9451	14586	9609	12839	10147	16914	21408	20674
1	P16	9618	12347	9878	19786	12445	9276	9389	12644	13003	8474	21719	12085	10333	11298	8758	12872	9140	12815	9283	15864	19808	17680
1	P17	7434	9402	6909	16186	9516	7184	6412	9146	10219	5189	17278	8866	6989	8738	6811	9458	6882	9119	7132	12697	15047	14129
1	P19	9729	11951	9287	18683	12009	8893	8606	11850	11499	7942	19457	11261	10176	10080	8012	11723	8652	11173	9205	14922	18674	17203
1	P20	10412	12685	9880	21286	13737	10789	9609	11830	13533	9948	21222	12217	11350	10989	9232	12751	9627	12647	9763	16282	21354	19428
1	P21	19484	25137	19417	42715	28358	20221	18489	23966	25717	18755	41008	23813	21876	20781	16951	24480	17823	23963	18761	32217	41647	39094
2	P22	13018	15941	11858	26206	19144	12581	12963	15561	16924	9327	30748	16518	12941	14664	11613	15370	11526	16388	11257	21637	25120	24744
1	P23	8737	10704	8403	17149	10997	8777	7916	10454	11365	8250	18093	10059	9561	9273	7241	10964	7785	10109	7956	13235	17236	15542
1	P24	9682	11750	9046	17831	12055	9116	8486	10951	11682	7882	19186	11172	9834	9607	8251	11607	8418	11035	8636	14776	17987	16414
1	P25	12132	15088	11415	27807	17158	12116	11206	15211	15434	9480	26054	14983	12803	12593	10442	14819	10464	14245	10943	19255	25610	23417
1	P26	6167	8211	6353	12999	8506	6651	6321	8090	8751	5995	13550	8389	7107	6893	5687	8294	5808	7551	5874	10937	14161	12577
1	P27	8057	10290	7916	16445	10383	7940	7545	10073	10266	6966	16252	9489	8908	7931	6578	10004	6712	9012	7225	12325	15705	14796
1	P28	8997	10960	8229	17612	11093	9063	8066	10478	11110	8237	17871	10541	9995	8643	7421	11024	7838	10068	8072	13598	17340	15747
1	P29	7820	9837	7372	15637	10103	7878	7116	9182	10133	6984	17193	9357	8327	8518	6634	9414	6489	9065	7455	12486	16366	15616
1	P30	7086	9434	6764	15757	9679	6986	6389	8904	9027	6081	15466	8774	7522	7770	6280	8755	6295	8162	6505	11568	15104	14096
1	P31	7381	9504	6830	14645	9562	7308	6916	8984	10007	6104	16834	9197	7342	8442	6906	9425	6783	9344	6306	11929	14774	14001
1	P32	6201	7856	6207	11992	8080	5883	5944	7299	7991	6180	13191	7980	7386	6720	5389	8392	5982	7808	5957	9829	12253	11332
1	P33	7483	9857	7233	14141	9817	7475	6962	8853	9881	6497	14718	9260	8316	7844	6472	9729	6898	8991	7125	11729	14891	12753
1	P34	8823	11347	8288	16813	11158	8539	8841	10347	11572	7930	18369	11501	10422	9188	7865	11383	7972	10815	8371	14333	17428	16455
1	P35	6278	8010	6048	12754	8049	6417	5741	7884	8133	4945	13780	8366	6375	6731	5430	8348	5495	7750	5979	9573	11142	11189
1	P36	8979	11257	8513	18054	11537	8441	8369	10736	11755	7428	18857	11201	9771	9979	7966	11399	8160	11130	8130	13617	16789	16078
1	P37	9212	11939	8990	17851	11858	9079	8719	11092	12087	8808	19384	11105	10524	9873	8189	12114	8693	11412	8342	14442	17920	16890
2	P38	8988	11000	7959	17999	12073	8839	8576	10688	11501	8084	18466	11103	9419	9603	7583	10519	7282	10679	7944	13270	14803	14997
1	P39	12828	16536	12284	27919	17211	13997	12236	15856	16263	10997	26991	15260	13510	13542	11380	16002	11474	14382	11643	19839	25640	24259
1	P40	6273	7887	5488	12492	7822	6788	5507	7203	8115	5103	12401	7374	6404	6357	5180	7624	5275	7060	5547	9298	11428	10998
1	P41	6273	7286	5888	11693	7882	6056	5767	7494	7956	5188	11863	7002	6498	6343	5122	7795	5328	7204	5513	9511	10940	10015
1	P42	8433	8090	6097	12328	8432	6258	5563	7892	8213	5332	12887	7471	7004	6719	5391	7825	5328	6883	5216	9395	11883	11550
3	ZIC3	5054	6358	4427	4881	6889	2737	4817	5804	3532	2301	9878	3183	2707	2733	2444	6183	2207	3058	2448	7146	8185	4283
1	P43	5030	6152	4737	9327	6291	5003	4557	5772	6520	4298	9057	5650	5637	5081	4328	6231	4288	5841	4650	7436	8486	8401
1	P44	5207	6077	4678	8836	6152	5033	4726	5708	6339	4271	9991	5743	5436	5245	4309	6193	4357	5740	4428	6972	8520	8233
1	P45	11072	13781	10386	21147	14881	11298	11133	14110	14198	9988	21887	13457	12078	11810	9733	14172	10058	12981	10255	16591	17983	17945
1	P46	8078	8431	5985	11388	7777	6280	5715	7436	7867	4720	12235	7344	6316	6424	5125	7530	5715	7138	5405	8883	10022	9545
4	SRY	0	0	0	3545	0	2323	0	0	2828	2291	0	2885	2485	2420	1982	0	2154	2833	2089	0	0	0
1	IRX3	2554	3075	2278	4222	3248	2388	2316	2845	3072	2291	4568	2934	2789	2851	2048	3104	2246	2868	2229	3502	3552	3328
1	nkx18	4054	5047	3591	7883	5388	3871	3838	4802	5225	3703	8038	4648	4412	4301	3388	4874	3547	4833	3983	6000	6710	6738
1	P47	3474	4675	3838	6488	4858	3888	3387	4357	4837	3056	6946	4461	3918	3979	3144	4552	3024	4001	3197	5187	6008	5768
1	MEF2C	1278	1531	1168	1778	1473	1291	979	1241	1345	1107	2081	1228	1287	1342	980	1177	824	1488	1134	1838	1524	1342
1	P48	2715	3014	2350	4498	3319	2754	2375	3147	3808	2125	4982	2981	2708	2585	2348	2940	2180	2884	2198	3857	3988	4027
1	P49	2718	3403	2560	4825	3542	2754	2788	3081	3738	2315	5328	3291	3078	2834	2535	3471	2182	3060	2308	3818	4345	4268

Figure 4.5

The Maphematica macro (John Armour) was the run in the new table. This macro made a new table in the same sheet replacing the values of the peak areas for a number resulting from the division of the area of each peak by the sum of the areas of the 4 nearest peaks in the same sample. The last column was the mean of the values in the entire row see Figure 4.6). A third table was constructed by the macro, this time replacing the values in the cells of the second table with the result of its value divided by the mean of the relevant row (normalized ratio). The last column in this table contained the standard deviations calculated with the values of the normalized ratios of each entire row, corresponding to the same probe in different samples (see Figure 4.7).

Ideally, the values in this last table represented, in the case of the autosomal probes, the normal diploid dosage as 1. In the case of the X-linked probe the normal female dosage was ideally represented as 1 and the normal male dosage as 0.5 whereas for the Y-linked probe the normal female value was 0 and the normal male value 1.

ratios relative to 4 neighbours

	2	3	6	7	11	12	13	15	17	18	19	20	21	23	24	25	27	28	29	30	31	32	mean
P01	0.3104034	0.3006375	0.3081582	0.3324602	0.3103175	0.3322957	0.3196963	0.3031262	0.2955659	0.2972179	0.3102163	0.2858122	0.3044299	0.2979191	0.3118494	0.2976237	0.298931	0.2947165	0.2842602	0.309608	0.3528055	0.3471499	0.3093273
P02	0.1913734	0.1852116	0.1657227	0.1659213	0.1789058	0.1720683	0.1811428	0.1865001	0.190804	0.1624196	0.1791086	0.1932722	0.168056	0.1648581	0.1859797	0.1761421	0.1827063	0.1768057	0.181228	0.1820978	0.1573647	0.1520931	0.1763537
P03	0.2412763	0.2499261	0.2418542	0.2297714	0.2326003	0.2365657	0.2376596	0.2379417	0.2446155	0.2312115	0.2438747	0.2438519	0.2417811	0.2462878	0.2399522	0.2467743	0.2412704	0.2411103	0.2327854	0.2437917	0.2333047	0.2215316	0.239079
P04	0.2854598	0.2820839	0.3111371	0.3027682	0.3095949	0.2830945	0.288592	0.2972989	0.2978038	0.2967913	0.2887607	0.2986211	0.2927855	0.2965796	0.2996329	0.3114506	0.2887262	0.3095238	0.3084643	0.2726417	0.2858764	0.3018644	0.2959796
P05	0.254902	0.248455	0.2344212	0.2606113	0.2477962	0.246699	0.239844	0.2398301	0.2389411	0.2581561	0.250076	0.2393172	0.2360924	0.2573812	0.2414067	0.2348588	0.2406616	0.2435524	0.250236	0.2589024	0.2630096	0.268796	0.2479066
P06	0.2152498	0.2211566	0.23173	0.2037016	0.2055153	0.2420291	0.2447785	0.2259056	0.2166285	0.2402723	0.2165729	0.224951	0.2431585	0.2188237	0.2193944	0.2302532	0.2302409	0.2225085	0.2288566	0.2254175	0.2120491	0.2067665	0.2239073
P07	0.2775932	0.2848885	0.2864038	0.3126077	0.2941636	0.2729148	0.2749158	0.2975955	0.2832811	0.2531011	0.2912162	0.2700567	0.2716168	0.2756242	0.2885369	0.261494	0.2794619	0.278021	0.2596349	0.2796174	0.2965842	0.2976708	0.2812274
P08	0.2247247	0.2332994	0.2136811	0.2370145	0.2479935	0.2238719	0.2147205	0.2256059	0.2382403	0.238982	0.2583907	0.2442986	0.2340414	0.2291562	0.2154806	0.2303817	0.2451377	0.2184237	0.2363335	0.240529	0.2479029	0.2541417	0.2341978
P09	0.2637777	0.246638	0.2576177	0.2223679	0.2433691	0.2612428	0.2553956	0.2387634	0.2521565	0.2737455	0.2174425	0.2538856	0.2706166	0.2588678	0.2630519	0.2719733	0.268438	0.2713125	0.2690874	0.2420049	0.2320899	0.2315626	0.2529731
P10	0.2610418	0.2568783	0.2598401	0.2420986	0.2381734	0.2525304	0.2631925	0.2526252	0.2472611	0.2518221	0.2418532	0.2393074	0.2509682	0.2547535	0.2551506	0.250416	0.2358072	0.2660127	0.2393441	0.2488596	0.2489079	0.2342003	0.2495929
P11	0.2905618	0.301707	0.315816	0.3319994	0.307974	0.2903068	0.3022286	0.3048047	0.3167338	0.2775533	0.33114	0.3177556	0.2782146	0.2936036	0.3015842	0.2977562	0.2796002	0.2701653	0.2950144	0.3103225	0.300337	0.3225813	0.3017164
P12	0.1603374	0.1644311	0.1565834	0.1585208	0.1647143	0.1651594	0.1663346	0.1710087	0.1635428	0.168632	0.1661032	0.1668603	0.1704704	0.16673	0.1679669	0.1662975	0.1629572	0.1735403	0.1751073	0.1631279	0.1603879	0.1525414	0.1650616
P13	0.1502549	0.150507	0.1510538	0.1337542	0.1420818	0.1570007	0.1552112	0.1519442	0.1461511	0.1420956	0.1342215	0.1449946	0.1471157	0.153471	0.1452962	0.1481099	0.1493126	0.1468044	0.1493925	0.1449023	0.140308	0.1350466	0.1463195
P14	0.4944974	0.4897679	0.4701908	0.538229	0.5227406	0.4910816	0.4682504	0.4776352	0.4867745	0.5286034	0.5081092	0.4975718	0.5031102	0.4898923	0.5000845	0.4682133	0.5182615	0.4890279	0.4847741	0.5150995	0.5502522	0.5485761	0.501852
P15	0.2740949	0.2675394	0.2875738	0.2670348	0.2680792	0.2810664	0.2696784	0.2641426	0.2699421	0.2969356	0.2618776	0.2625641	0.2929163	0.2463057	0.2634499	0.2871316	0.257538	0.2591066	0.2686381	0.2570986	0.2558256	0.2724386	0.2695899
P16	0.2178828	0.2242381	0.2269274	0.2122438	0.2154754	0.2128157	0.238754	0.2343521	0.2270196	0.208278	0.2285513	0.2288306	0.2112742	0.2375205	0.2241445	0.2289903	0.2242835	0.240436	0.2216677	0.220821	0.21607	0.2076165	0.2230997
P17	0.1832705	0.1867366	0.1733447	0.1970442	0.1824421	0.1793355	0.1715172	0.1851492	0.1955228	0.1390712	0.203027	0.183679	0.1567216	0.2014943	0.1921135	0.1821228	0.1858593	0.184319	0.1857388	0.1984464	0.1852075	0.1884244	0.1836631
P19	0.2072293	0.2006178	0.2023576	0.1868805	0.1935188	0.1873394	0.1960865	0.2057077	0.1840665	0.1875502	0.1922116	0.1976273	0.2013136	0.1945796	0.191895	0.1968234	0.1990247	0.1908479	0.2048332	0.1936413	0.1908314	0.190444	0.1952467
P20	0.2250513	0.2155956	0.2183657	0.2186094	0.2277052	0.2367359	0.2239965	0.2052893	0.2239154	0.2466039	0.2133679	0.2180634	0.2298781	0.2159151	0.2277707	0.2178429	0.2265336	0.221605	0.2199815	0.2150859	0.2243633	0.2205071	0.2223992
P21	0.5365719	0.5618211	0.5629911	0.5827104	0.5697918	0.5673204	0.5681057	0.5542052	0.5516775	0.599029	0.5392242	0.5615876	0.5745351	0.5317554	0.5416347	0.5452602	0.5409761	0.5566577	0.5508868	0.5638652	0.5759428	0.5896353	0.5602811
P22	0.2683294	0.2644668	0.2536145	0.2653142	0.3031656	0.2572644	0.2914344	0.2719456	0.2716664	0.2080294	0.3089972	0.2884686	0.2459284	0.2895163	0.2786563	0.2570148	0.2640368	0.2837552	0.2495124	0.2827996	0.2557941	0.2734808	0.2696905
P23	0.1683105	0.1655992	0.1688434	0.1569861	0.1586686	0.1680066	0.1657107	0.1686728	0.1712473	0.1790948	0.168354	0.1617593	0.1711509	0.1718177	0.1613557	0.1722356	0.1680264	0.1633382	0.1653951	0.1603659	0.1617219	0.1580226	0.166122
P24	0.2124248	0.198755	0.1984294	0.1754852	0.1911096	0.190851	0.1927069	0.1896573	0.1906736	0.1855461	0.1943772	0.1951646	0.1915204	0.1939241	0.2046328	0.1982171	0.2010125	0.1975192	0.1983737	0.195336	0.1821214	0.18111	0.1935885
P25	0.3693938	0.3678672	0.3599584	0.4298549	0.4085141	0.3729836	0.3704708	0.3844268	0.3669171	0.325818	0.3883961	0.3831087	0.3615645	0.3736352	0.3761934	0.3625976	0.3643327	0.3777813	0.3685629	0.3755388	0.3935822	0.3946974	0.3761907
P26	0.1590745	0.1708276	0.1735792	0.1639322	0.1689913	0.1739506	0.1791413	0.1731852	0.1804628	0.1840764	0.1705196	0.1816391	0.1723327	0.1777738	0.1739569	0.1747798	0.1736659	0.1702209	0.1684253	0.1824232	0.1848164	0.1787166	0.174386
P27	0.2307538	0.2345353	0.2370464	0.2217652	0.2212916	0.2236116	0.2306564	0.2345777	0.225984	0.2291982	0.2173657	0.2192974	0.2348413	0.2164161	0.21793	0.2297077	0.2193679	0.2201862	0.2233799	0.2190021	0.2136822	0.2196654	0.2245574
P28	0.2975164	0.2917066	0.2897637	0.2880696	0.2864114	0.3130029	0.2947645	0.2892158	0.2910129	0.3202691	0.2877155	0.2927324	0.3042619	0.2776028	0.2947297	0.3023007	0.3097779	0.297958	0.2983111	0.2873748	0.2826129	0.2756518	0.2942347
P29	0.2536592	0.2397979	0.2471089	0.2472424	0.2481273	0.2477174	0.2460751	0.2384758	0.2507548	0.2445654	0.258452	0.2462304	0.2488123	0.259806	0.2440316	0.2401041	0.2348704	0.2477724	0.265227	0.2526912	0.2604135	0.2663029	0.2494654
P30	0.2347098	0.2485444	0.2353678	0.2631166	0.2492147	0.2233133	0.2278206	0.2480707	0.23004	0.2228719	0.2372486	0.2367156	0.2296654	0.2403661	0.2383302	0.228859	0.2323564	0.2249538	0.234077	0.2417858	0.2486132	0.2486242	0.2375766
P31	0.2574289	0.2597857	0.2513055	0.2545761	0.2606941	0.2634502	0.2615635	0.262044	0.2703718	0.2402015	0.2779355	0.2623891	0.2326879	0.2736289	0.2787487	0.2597134	0.2643002	0.2746297	0.2331928	0.2615206	0.2519698	0.2602561	0.2596543
P32	0.2029455	0.1966852	0.2126049	0.19056	0.2060331	0.1963941	0.2039948	0.1966386	0.1974695	0.2316486	0.2016448	0.1977693	0.2198679	0.2021539	0.1972331	0.2135804	0.2140404	0.2092088	0.2104426	0.1983293	0.1970094	0.1977489	0.2042729
P33	0.2620067	0.2630117	0.2634685	0.2529515	0.2392673	0.2654946	0.2547919	0.2582254	0.2615442	0.2584431	0.2366066	0.252158	0.2637739	0.2523691	0.2549035	0.2591063	0.2629613	0.251743	0.2677263	0.2568544	0.2678478	0.2407271	0.2566366
P34	0.2981467	0.3085101	0.2952752	0.2968344	0.3056885	0.3025225	0.3268875	0.2993577	0.3066243	0.3168198	0.3048918	0.3152168	0.3272214	0.293811	0.3034802	0.3005968	0.3004334	0.3031365	0.3078592	0.3188796	0.3164049	0.3204354	0.3076915
P35	0.1831841	0.1812217	0.1849315	0.1904748	0.1851537	0.191358	0.1743872	0.1872776	0.1797149	0.1612693	0.1936536	0.1942322	0.1633233	0.1825455	0.1792552	0.18707	0.1732182	0.1830075	0.1870308	0.1762302	0.166234	0.1799569	0.1811241
P36	0.2843525	0.2748761	0.2815239	0.292804	0.2885404	0.2678832	0.2762684	0.2827575	0.2823481	0.2635912	0.2814834	0.2784171	0.2741813	0.2964262	0.287001	0.2741858	0.2808177	0.285619	0.2726632	0.2759151	0.27353	0.280657	0.279811
P37	0.2509466	0.2532024	0.2474867	0.2385832	0.2493744	0.2454182	0.2477904	0.2485491	0.252855	0.2850485	0.2490652	0.2396056	0.262588	0.2503995	0.2525832	0.2570228	0.2626205	0.2589105	0.2444685	0.2508947	0.252405	0.2484518	0.2521941
P38	0.240266	0.2317178	0.2274975	0.2325267	0.2482675	0.2396367	0.2462174	0.2381101	0.23861	0.190506	0.2388904	0.2469528	0.209381	0.2491822	0.2318606	0.2231486	0.2170109	0.2427928	0.2359931	0.2312008	0.206236	0.2207941	0.2312181
P39	0.4173472	0.4311487	0.4296757	0.4588636	0.4378275	0.463334	0.4314223	0.4341446	0.4076655	0.3998114	0.4330666	0.4098493	0.4069645	0.4161254	0.4292137	0.4110244	0.417905	0.3907515	0.4228897	0.4214966	0.4492177	0.4519272	0.4259852
P40	0.1805388	0.179358	0.1668133	0.1795266	0.1741329	0.1648377	0.1705746	0.1709465	0.1823636	0.1706518	0.174306	0.1778668	0.1706095	0.1742742	0.1726432	0.1743186	0.1711495	0.1770267	0.1806017	0.1748172	0.1722044	0.170552	0.1740961
P41	0.2052415	0.1887596	0.205759	0.1893358	0.1980128	0.1978058	0.206977	0.2051745	0.203421	0.203964	0.1937361	0.2113134	0.1996007	0.2001009	0.1949827	0.2068627	0.2022549	0.2108529	0.2037626	0.2068958	0.1905359	0.1824027	0.2003524
P42	0.211584	0.2137271	0.2146302	0.2015344	0.2144674	0.2056794	0.1982043	0.2117433	0.2113811	0.2119068	0.2136014	0.2069185	0.2165404	0.2145069	0.2073462	0.2078243	0.2021636	0.1995825	0.1906921	0.203858	0.2099869	0.2164097	0.2084677
ZIC3	0.2202851	0.2302455	0.2068692	0.1109873	0.232871	0.1224499	0.2239849	0.2176718	0.1216756	0.1206924	0.2223448	0.1193655	0.1101436	0.1168548	0.127624	0.220475	0.1144769	0.1191367	0.1235927	0.2144915	0.2051044	0.1121234	0.2189979
P43	0.173538	0.1746586	0.1751266	0.172735	0.1679526	0.1746553	0.1676045	0.1649049	0.1776276	0.1757918	0.1607333	0.1706883	0.1817449	0.1687067	0.1762574	0.1731555	0.1702637	0.1781444	0.1829844	0.1751395	0.1721089	0.175963	0.173204
P44	0.1819802	0.1667948	0.1718905	0.1630469	0.1641365	0.1745327	0.1752447	0.162982	0.1732062	0.1778324	0.1728499	0.1683078	0.1752328	0.1746354	0.1753265	0.173192	0.1717721	0.174888	0.1734702	0.1648734	0.1762005	0.1731768	0.1720715
P45	0.5667826	0.5797767	0.5871777	0.6259472	0.6383586	0.6036768	0.6430057	0.6484673	0.6016612	0.620683	0.6156508	0.6152897	0.5981179	0.6087315	0.6157009	0.614624	0.6062945	0.6004633	0.6136309	0.6196452	0.5880641	0.6061067	0.6108116
P46	0.2547039	0.290074	0.271525	0.2618533	0.2535537	0.2646662	0.2514077	0.2615271	0.2544724	0.2299411	0.2706679	0.2624357	0.2433911	0.2591681	0.2510286	0.2535354	0.2730922	0.2603429	0.2506725	0.2568762	0.2600348	0.2544385	0.2586095
SRY	0	0	0	0.0797507	0	0.097429	0	0	0.0970095	0.1123976	0	0.1019977	0.0974495	0.0960851	0.0976355	0	0.0998887	0.1026524	0.0958669	0	0	0	0.0980148
IRX3	0.103493	0.0963527	0.0964927	0.0903546	0.0990123	0.0953328	0.0965604	0.0926559	0.0962195	0.1083369	0.0930384	0.0980943	0.1047452	0.0999849	0.0956164	0.0997173	0.1005282	0.0998225	0.098847	0.0956282	0.0872277	0.0830008	0.0968664
new18	0.3030801	0.2849481	0.2746463	0.3216125	0.3142124	0.2854088	0.3109523	0.302412	0.3087879	0.331424	0.3111197	0.2911369	0.3078856	0.2987635	0.3038926	0.2978671	0.297842	0.2990383	0.3086502	0.3129727	0.3179492	0.3355076	0.3054595
P47	0.3279524	0.3690692	0.3874267	0.3583324	0.3466865	0.3605948	0.3481053	0.3620274	0.3650566	0.3311294	0.353648	0.3784036	0.3493224	0.3664242	0.3634262	0.3763539	0.3398899	0.3366995	0.3455469	0.3491658	0.3812849	0.3736961	0.3577384
MEF2C	0.0980468	0.0948634	0.0960699	0.0754315	0.0871237	0.091925	0.0796453	0.0806525	0.0772633	0.0988569	0.0815076	0.0797087	0.0918816	0.0981066	0.0752801	0.0743196	0.0844375	0.1017923	0.0995435	0.0879292	0.0724782	0.0645223	0.0859721
P48	0.2359228	0.2056496	0.2145531	0.2146226	0.2203705	0.2392286	0.2179699	0.2334397	0.2382304	0.2087426	0.2227588	0.2187729	0.2131444	0.2059249	0.2364368	0.2068958	0.2250439	0.2190433	0.2125847	0.2200096	0.2133979	0.2223142	0.2203208
P49	0.2353861	0.2385225	0.2382947	0.2364037	0.2387114	0.2392286	0.264989	0.2274304	0.248951	0.2317317	0.2419728	0.2471463	0.2495338	0.2325429	0.2600267	0.2562948	0.2265633	0.2374001	0.225636	0.2317925	0.2386576	0.2386573	0.240267

Figure 4.6

normalise ratios relative to means

	2	3	6	7	11	12	13	15	17	18	19	20	21	23	24	25	27	28	29	30	31	32	STD DEV
P01	1.0034788	0.9719075	0.9962206	1.0747845	1.0032012	1.0742529	1.0335212	0.9799531	0.9555117	0.9608525	1.0028739	0.92398	0.9841676	0.9631192	1.0081534	0.9621642	0.9663906	0.9527661	0.9189625	1.0009075	1.1405572	1.1222737	0.0579414
P02	1.0851679	1.050228	0.9397174	0.9406437	1.0144716	0.9756999	1.0271563	1.0575343	1.0819394	0.9209874	1.0156214	1.0959349	0.9529482	0.9348148	1.0545835	0.9987998	1.0360216	1.0025629	1.0276393	1.0325714	0.8923244	0.9624318	0.0644589
P03	1.0091906	1.0453703	1.0116078	0.9610687	0.9729015	0.9894876	0.9940628	0.9952429	1.0231577	0.9670923	1.0200588	1.0199636	1.0113019	1.0301524	1.0036525	1.0321671	1.0091661	1.0084962	0.9736754	1.019712	0.9758476	0.9266042	0.0279544
P04	0.9644577	0.9530519	1.0512113	1.0229359	1.0460007	0.9564662	0.9750399	1.0044573	1.0061633	1.0027425	0.97561	1.0089244	0.9892084	1.0087844	1.0123429	1.0522706	0.9754936	1.0457605	1.0421808	0.9211501	0.9658652	1.0198825	0.0360687
P05	1.0282175	1.0022121	0.9456026	1.0512479	0.9995544	0.9951285	0.9674771	0.967421	0.963835	1.0413441	1.0087508	0.9653519	0.952344	1.0382182	0.9737806	0.9473679	0.9707749	0.9824359	1.0093961	1.0443545	1.0609218	1.0842631	0.0402629
P06	0.9613344	0.9877151	1.0349372	0.9097586	0.9178591	1.0809345	1.0932138	1.0089248	0.9674918	1.0730886	0.9672437	1.0046616	1.0859787	0.9772959	0.9798448	1.0283417	1.0282869	0.9937528	1.0221042	1.0067449	0.9470399	0.9234471	0.0533287
P07	0.9870775	1.0130182	1.0184085	1.1115833	1.0459991	0.9704415	0.9775677	1.0582022	1.0073027	0.8999873	1.0355186	0.9602788	0.9658252	0.9800759	1.0259915	0.9298309	0.993722	0.9885984	0.9232204	0.9942749	1.0546063	1.0584699	0.0500303
P08	0.9595509	0.9961639	0.9123958	1.012027	1.0589061	0.9559095	0.916834	0.9633133	1.0172611	1.0204282	1.103301	1.0431295	0.9993322	0.9784728	0.9200798	0.9837058	1.0467122	0.9326461	1.0091193	1.0270336	1.0585194	1.0851586	0.0542124
P09	1.0427107	0.9749577	1.0183603	0.8790182	0.9620355	1.0326902	1.0095761	0.9438293	0.9967721	1.0821133	0.859548	1.0036073	1.0697449	1.023302	1.0398417	1.0751076	1.0611329	1.0724956	1.0637	0.9566429	0.9174489	0.9153647	0.0659071
P10	1.0458703	1.029189	1.0410556	0.9699739	0.9542474	1.011769	1.0544871	1.0121488	0.9906574	1.0089312	0.9689907	0.9587908	1.0055101	1.0206762	1.0222668	1.0032975	0.9447673	1.0657862	0.9589381	0.9970618	0.9972556	0.938329	0.0382313
P11	0.9630297	0.9999688	1.0467314	1.1003692	1.02074	0.9621843	1.0016978	1.0102359	1.0497732	0.9199146	1.0975208	1.0531601	0.9221054	0.9731113	0.999562	0.9868744	0.9266989	0.8954279	0.977787	1.028524	0.9954282	1.0691541	0.0561466
P12	0.9713795	0.9961804	0.9486365	0.9603737	0.997896	1.0005924	1.0077122	1.0360295	0.9907965	1.0216311	1.0063104	1.0108971	1.0327687	1.0101077	1.0176011	1.0074877	0.9872511	1.0513671	1.0608607	0.9882848	0.9716852	0.9241487	0.0322601
P13	1.0268958	1.0286185	1.0323557	0.9141243	0.971038	1.072999	1.060769	1.0384411	0.9988488	0.9711322	0.9173177	0.9909446	1.0054413	1.0488755	0.9930063	1.0122358	1.0204556	1.0033135	1.0210015	0.9903144	0.9589149	0.9229567	0.0437853
P14	0.9853452	0.9759211	0.9369114	1.0724856	1.0416231	0.9785387	0.9330448	0.9517453	0.9699564	1.0533054	1.0124683	0.9914712	1.0025072	0.9761689	0.9964781	0.9329708	1.032698	0.9744465	0.9659702	1.0263973	1.0964432	1.0931033	0.048615
P15	1.0167106	0.9923939	1.0667085	0.9905223	0.9943963	1.0425703	1.0003283	0.9797939	1.0013065	1.1014342	0.9713923	0.9739387	1.0865254	0.9136312	0.9772244	1.0650682	0.9552953	0.9611145	0.9964696	0.9536653	0.9489436	1.0105666	0.0473107
P16	0.9766163	1.0051029	1.017157	0.9513405	0.9658259	0.9539043	1.0701672	1.0504368	1.0175701	0.9335649	1.0244357	1.0256876	0.9469947	1.0646384	1.0046834	1.0264033	1.0053061	1.0777067	0.9935814	0.9897862	0.9684908	0.9305998	0.0434808
P17	0.9978623	1.0167345	0.943819	1.072857	0.9933523	0.9764373	0.9338689	1.0080916	1.0645733	0.7572065	1.1054315	1.0000865	0.8533104	1.0970864	1.0460105	0.9916134	1.0119582	1.0035715	1.011302	1.0804917	1.0084091	1.025924	0.078272
P19	1.0613715	1.0275091	1.0364202	0.9571505	0.99115	0.9595009	1.0043012	1.0535786	0.942738	0.9605806	0.9844549	1.0121927	1.0310731	0.9965832	0.9828336	1.0080756	1.0193497	0.9774707	1.0490996	0.9917778	0.9773863	0.9754021	0.0333313
P20	1.011925	0.9694082	0.9818635	0.9829595	1.023858	1.0644636	1.0071819	0.9230666	1.0068175	1.1088343	0.9593915	0.9805042	1.0336281	0.9708448	1.0241522	0.979513	1.0185901	0.9964291	0.9891291	0.9871161	1.0088313	0.9914923	0.0384244
P21	0.9576834	1.0027486	1.0048388	1.0400321	1.0169748	1.0125637	1.0139654	0.9891555	0.9846441	1.0691579	0.9624172	1.0023318	1.0254408	0.9490867	0.9667195	0.9731903	0.9655441	0.9935328	0.9832328	1.0063968	1.0279532	1.0523918	0.0314435
P22	0.994953	0.9806307	0.9403909	0.9837728	1.1241242	0.9539245	1.0806252	1.0083618	1.0073264	0.7713636	1.1457473	1.0696283	0.9118913	1.0735131	1.0332446	0.9529991	0.9790364	1.0521514	0.9251804	1.048608	0.9484726	1.0140544	0.0608261
P23	1.0131745	0.9968531	1.0163824	0.9450048	0.9551329	1.0113448	0.9975243	1.015355	1.030853	1.078092	1.0134359	0.973738	1.0302723	1.0342864	0.9713089	1.036802	1.0114642	0.9832426	0.9956245	0.9853506	0.9735133	0.9512447	0.0329217
P24	1.0973003	1.026688	1.025006	0.9064856	0.9871946	0.9858592	0.9954459	0.9796926	0.9849426	0.9584562	1.0040738	1.0081411	0.989317	1.0017334	1.0570503	1.0239094	1.038349	1.0203039	1.0247181	1.0090269	0.9407652	0.935541	0.0417276
P25	0.9819323	0.9778742	0.9568508	1.1426515	1.0859229	0.9914749	0.9847952	1.0218935	0.9753486	0.8660979	1.0324448	1.0183897	0.9611203	0.9932068	1.0000072	0.9638664	0.9684788	1.0042282	0.9797235	0.9982671	1.0462305	1.0491949	0.053438
P26	0.9121978	0.979595	0.9953739	0.9400536	0.9690648	0.9975033	1.0272689	0.9931143	1.0348468	1.0555689	0.9778288	1.0415922	0.9882257	1.0194271	0.9975398	1.0022584	0.9658712	0.976116	0.965819	1.0460888	1.0598122	1.0248335	0.0389188
P27	1.0275941	1.0444336	1.0556161	0.987566	0.9854567	0.9957883	1.0271601	1.0446225	1.006353	1.0206665	0.9679741	0.9765765	1.0457965	0.9637452	0.9704871	1.0229355	0.9768903	0.9805342	0.9947566	0.9752613	0.9515706	0.9782149	0.0309874
P28	1.0111536	0.991408	0.9848049	0.9790472	0.9734115	1.0637868	1.0018009	0.9829427	0.9890504	1.0884818	0.9778438	0.9948943	1.0340789	0.9441538	1.0016826	1.0274137	1.0528259	1.0126542	1.0138544	0.9766857	0.9605019	0.9375231	0.0368863
P29	1.0168112	0.9612475	0.990554	0.9910892	0.9946364	0.9929931	0.9864099	0.9559474	1.0051687	0.9803582	1.0360235	0.9870323	0.997382	1.0414513	0.9782185	0.9624746	0.9414951	0.9932135	1.0631816	1.0129311	1.0438865	1.0674944	0.0336309
P30	0.987933	1.0461653	0.9907028	1.1075021	1.0489867	0.9399633	0.9589352	1.0441712	0.9682772	0.9381052	0.9986195	1.0047952	0.9667003	1.0118257	1.0031719	0.963306	0.9780274	0.9468687	0.9852696	1.0177172	1.0464551	1.0465014	0.0431972
P31	0.9914293	1.0005062	0.9678465	0.9804426	1.0040048	1.014619	1.0073529	1.0092034	1.041276	0.9250819	1.0704062	1.0105325	0.8961449	1.0538203	1.073538	1.0002278	1.0178927	1.0576744	0.8980896	1.0071877	0.9704052	1.002318	0.0479575
P32	0.993502	0.9628552	1.0407887	0.9328699	1.0086171	0.96143	0.9986387	0.9626272	0.9666949	1.1340156	0.9871347	0.9681623	1.076344	0.9896268	0.9655374	1.0455641	1.0478159	1.0241636	1.0302036	0.9709035	0.9644424	0.9680624	0.0472654
P33	1.0209249	1.0248407	1.0266206	0.9856406	0.9323193	1.0345155	0.9928119	1.0061907	1.0191225	1.0070388	0.9219594	0.9825489	1.0278108	0.9834491	0.9932467	1.0096311	1.0246443	0.9809315	1.0432115	1.0008486	1.043685	0.9380075	0.0340992
P34	0.9689794	1.0026602	0.9596467	0.9647144	0.9940751	0.9832006	1.0623861	0.972915	0.9965314	1.029667	0.9909008	1.0244572	1.0634721	0.9548881	0.9863132	0.9769421	0.976411	0.985196	1.000545	1.0363613	1.0283185	1.0414178	0.0321878
P35	1.0113732	1.000539	1.021021	1.0516258	1.0222475	1.0565023	0.9628049	1.033974	0.99222	0.89038	1.0691761	1.0723711	0.9017207	1.0078476	0.9896819	1.0328279	0.9563508	1.0103982	1.0326112	0.9729803	0.9177906	0.9935558	0.0502722
P36	1.0162307	0.9823633	1.0061214	1.0464348	1.0311975	0.9573719	0.9873394	1.0105304	1.0090673	0.942033	1.0059768	0.9950185	0.9798804	1.0593802	1.0256959	0.9798963	1.0035976	1.0207567	0.974455	0.9860766	0.9775528	1.0030236	0.0277155
P37	0.9950536	1.0039982	0.9813342	0.9460301	0.9888195	0.9731322	0.9825384	0.9855468	1.0026207	1.1302745	0.9875935	0.9500641	1.0412138	0.9928839	1.0015429	1.0191469	1.0413427	1.0266321	0.9693667	0.9948476	1.0008364	0.985161	0.0378019
P38	1.0391313	1.0021611	0.9839085	1.0056593	1.0737371	1.0364096	1.0648708	1.0296071	1.031969	0.8239233	1.0331819	1.0680514	0.9055561	1.0776932	1.0027785	0.9650997	0.9385548	1.0500596	1.0206514	0.9999249	0.8919542	0.954917	0.0647484
P39	0.9797222	1.0121211	1.0086635	1.077182	1.0277998	1.0876762	1.0127636	1.0191541	0.9569945	0.9385569	1.0166282	0.9621208	0.9553488	0.9768563	1.0075789	0.9648794	0.9810317	0.9172889	0.9927333	0.9894629	1.0545383	1.0608988	0.0440693
P40	1.0370066	1.0302242	0.9581682	1.0311928	1.0002113	0.9468202	0.9797727	0.9819088	1.0474886	0.9802161	1.0012058	1.0216587	0.9799736	1.0010232	0.9916548	1.0012785	0.9630753	1.0168332	1.037368	1.0041419	0.9891343	0.9796432	0.0264366
P41	1.0244024	0.942139	1.0269856	0.9450139	0.9883229	0.9872894	1.0330648	1.0240684	1.0153162	1.0180263	0.966977	1.0547087	0.9962481	0.998745	0.9731987	1.0324943	1.0094957	1.0524102	1.0170209	1.0326595	0.9510039	0.9104092	0.0382467
P42	1.0149488	1.0252293	1.0295612	0.9667419	1.0287802	0.9866248	0.9507676	1.0157131	1.0139754	1.0164975	1.0246261	0.9925689	1.0483178	1.0289698	0.9946203	0.9969138	0.96976	0.9573785	0.9147321	0.9778879	1.0072877	1.0380971	0.0329895
ZIC3	1.0058774	1.0513594	0.944617	0.5067961	1.063348	0.5591372	1.0227716	0.9939444	0.5556017	0.5511119	1.0152826	0.545053	0.5029437	0.5335886	0.5827636	1.0067446	0.5227305	0.5440082	0.5643554	0.9794227	0.9365585	0.5119836	0.3492143
P43	1.0019289	1.0083985	1.0111006	0.9972926	0.969581	1.0083792	0.9676716	0.9520849	1.0255403	1.014941	0.9280002	0.9854758	1.0493115	0.9740465	1.0176291	0.9997202	0.9830243	1.0285241	1.0564679	1.011175	0.9936775	1.0159293	0.0300244
P44	1.0575851	0.9693341	0.9989483	0.9475533	0.9538856	1.0143036	1.0184415	0.9471762	1.0065944	1.0334797	1.0045236	0.9781275	1.0183723	1.0149004	1.0189169	1.0065121	0.9982603	1.0163685	1.0081287	0.9561679	1.0239961	1.0064239	0.0291113
P45	0.9606604	0.9491907	0.9613074	1.0247794	1.045099	0.9883191	1.052707	1.0616486	0.9850191	1.0161611	1.0079224	1.0073313	0.9792182	0.9965945	1.0080045	1.0062415	0.9926046	0.983058	1.0046156	1.014462	0.9627585	0.9922972	0.0294039
P46	0.9848979	1.121668	1.0499421	1.0125435	0.9804503	1.0234204	0.972152	1.011282	0.9840028	0.889144	1.0466279	1.0147953	0.9411532	1.0021602	0.9706861	0.9803792	1.0560022	1.006703	0.969309	0.9932978	1.0055114	0.9838715	0.0455781
SRY	0	0	0	0.8136603	0	0.9340237	0	0	0.9897434	1.1467412	0	1.0406356	0.9942329	0.9803126	0.99613	0	1.0191188	1.0473151	0.9780864	0	0	0	0.7257695
IRX3	1.0684092	0.9946964	0.9961419	0.9327752	1.0221529	0.9841678	0.9968401	0.9565327	0.9933213	1.118415	0.9604814	1.0126756	1.0813359	1.0321936	0.9870954	1.0294308	1.0378017	1.0305168	1.0204464	0.987217	0.9004948	0.8568582	0.0576067
new18	0.9922104	0.9328504	0.8991249	1.0528807	1.0286548	0.9343589	1.0179821	0.9900231	1.0106962	1.0850012	1.0185301	0.953111	1.0079422	0.9780789	0.9948703	0.9751443	0.9750619	0.9789783	1.0104455	1.0245963	1.0408881	1.0983702	0.0474566
P47	0.9167381	1.0316735	1.0629889	1.0016604	0.9691062	1.0079845	0.9730721	1.0119892	1.0204568	0.9256188	0.9885659	1.0577862	0.9764744	1.0242796	1.0158993	1.0520365	0.9501072	0.9411891	0.9659206	0.976045	1.0658204	1.0446073	0.0458866
MEF2C	1.1404493	1.1034208	1.1174544	0.8773953	1.0133951	1.0892425	0.926409	0.9381244	0.8987026	1.1498727	0.9480701	0.9271468	1.0687372	1.1411449	0.8756345	0.8644625	0.9821508	1.184016	1.157859	1.0227645	0.843044	0.7505034	0.1232845
P48	1.0708152	0.9334097	0.9738214	0.9741371	1.0002256	1.0858196	0.9693297	1.0595444	1.081289	0.9474487	1.0110657	0.9929746	0.9674277	0.9346592	1.0731482	0.948144	1.0214373	0.9942017	0.9648872	0.9965877	0.9685782	1.0090479	0.0475777
P49	0.9796856	0.9927393	0.9917913	0.983921	0.9935257	0.9956784	1.1028941	0.9465738	1.0361434	0.964476	1.0070999	1.029632	1.0385691	0.967852	1.0822406	1.0667083	0.9429648	0.9880681	0.9391054	0.9647289	0.9933017	0.9933007	0.0435911

Figure 4.7

4.4 Results

MAPH electrophoresis traces were obtained from the ABI3100 Genetic Analyzer corresponding to 40 DNA samples from control subjects. The control probe HEY2 showed an extremely low signal in control individuals due to its low GC content (30%), for that reason it was ignored during the analysis process. The signal from the X and Y chromosome probes displayed expected the variation pattern according to the sex of the control individuals (see Figure 4.8). The signals corresponding to probes 22 and 38 had considerable variation between samples from control subjects. This variability correlated strongly with the size of the batch of samples processed during the same experiment and it was probably due to the repetitive nature of those probes.

The analysis of the 380 samples from our cohort was performed in batches of variable size. The traces were inspected visually for a gross assessment of the quality of the data. The .fsa files generated by the ABI3700 Genetic Analyzer were processed, obtaining Maphematica Excel tables for each batch.

For autosomal probes with no potential cross hybridization targets, a range between 0.7 and 1.3 was established to consider a normalized ratio as "normal". For autosomal probes with one potential cross hybridization

target the normal range was established as 0.8 to 1.2. When a particular sample showed a normalized ratio outside the relevant ranges, a detailed inspection of the electrophoresis trace was carried out and the assay repeated in order to confirm the potential copy number variation.

After the first round of hybridizations and analysis, most of the samples showed no probe with signal outside the reference intervals while some samples showed just one or few normalized ratios outside the reference ranges. Few other samples displayed abnormal normalized ratios in most of the probes, probably due to a low quality trace, (see Figure 4.9) and some produced a trace that did not allowed analysis by Genotyper because of the low intensity signal of the whole trace. New filters were prepared, hybridized and processed for those samples with at least one potential copy number variation and those with signal not strong enough to be analyzed on the first round. A third round of MAPH was required for few samples that showed a weak trace in the first round and potential abnormality in the second.

The trace from sample x204 showed an X-linked probe peak compatible with two copies, whereas the Y-linked probe peak appeared to indicate one copy. As this sample corresponds to a male patient, the finding could mean a duplication of ZIC3 (the gene to which the X-linked probe binds) or a partial or complete disomy of chromosome (Figure 4.10).

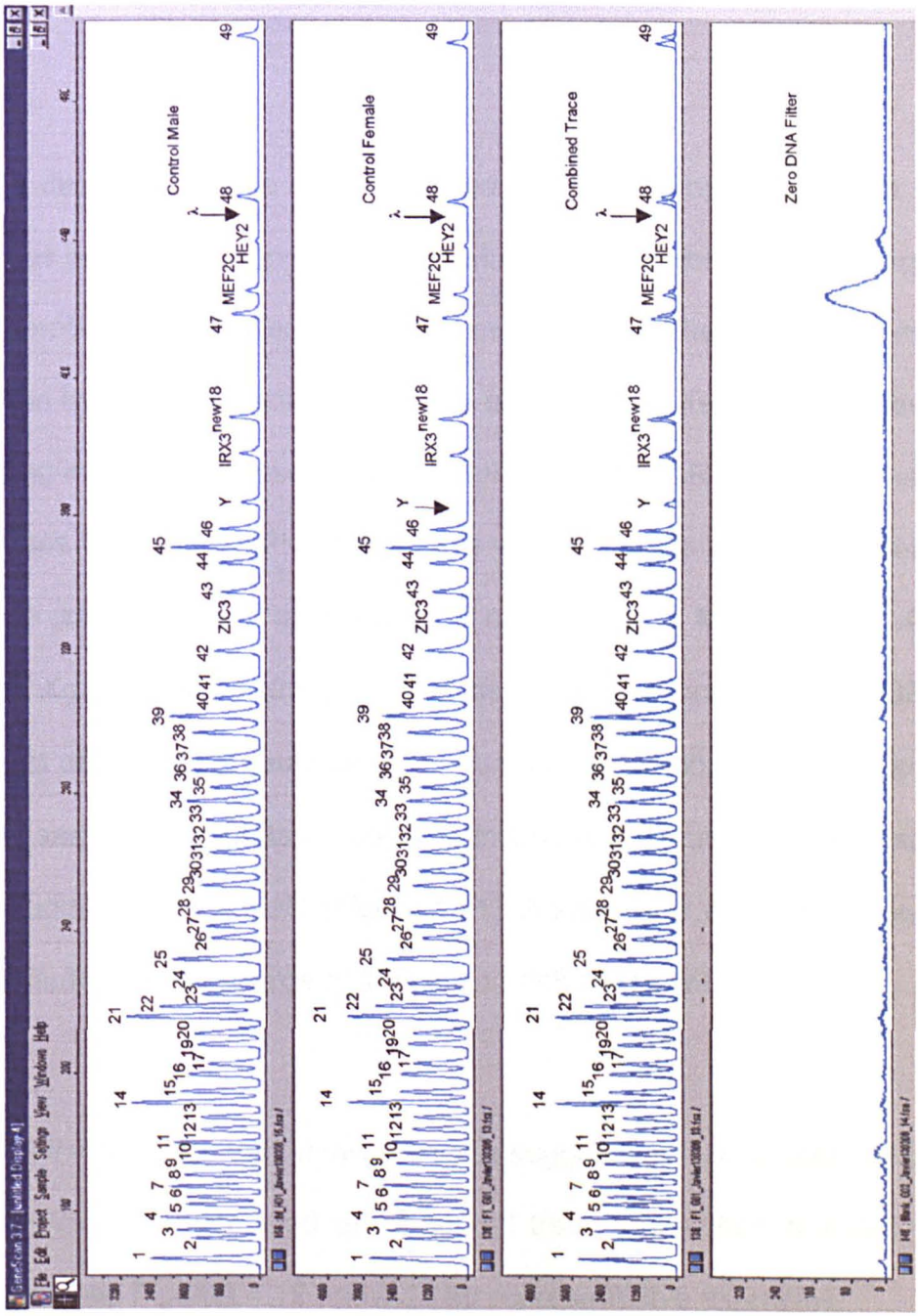


Figure 4.8: Two of the electrophoresis traces from control DNA used for the assessment of the MYH6 MAPH probe set. The expected variation in Y- and X- linked (ZIC3) probes according with sex was observed. The peaks from the HEY2 probe were too small for analysis and were ignored. The peak from the non-human (λ) probe was not observed, as expected. At least one zero DNA filter (bottom panel) was included in the hybridization in every batch as a negative control.

As the described profile was confirmed with a new hybridization, a new filter was processed along with filters with DNA from the parents using the XpYp probe set (provided by John Armour) that contains several probes that bind the X chromosome in several places distributed across its length including including the pseudoautosomal region 1 (PAR1). The analysis of the traces from the MAPH experiments with the XpYp probe set revealed that the patient carries an extra copy of the targets for probes in Xp21 (DMD), Xq25 (ZIC3), Xq28 (2d2) and the pseudoautosomal probe PGPL2. The rest of the pseudoautosomal probes showed a dosage similar to the control samples, indicating a possible incomplete disomy of chromosome X, excluding most of PAR1 (Figure 4.11). A karyotype was recommended to the cardiologist in charge of the case to define the structural defect.

None of the samples that showed a trace suggestive of an autosomal copy number variation displayed an abnormal trace when their analysis was repeated (see Figures 4.12 and 4.13 for representative examples).

It was concluded that copy number variation of the *MYH6* gene segments covered by the probes is not a significant contributor factor for Congenital Heart Defect in our patient cohort.

		1	2	3	4	5
	P1	1.001645	0.994519	1.010071	0.875276	1.560836
	P2	0.976996	0.995273	1.178115	1.177864	0.689621
	P3	0.949178	0.989444	0.93584	1.056349	0.738453
	P4	0.983586	0.957326	1.03539	1.041403	1.250844
	P5	1.036119	1.025284	0.868377	0.828857	0.709701
	P6	1.031476	1.026792	0.999861	1.103053	1.209428
	P7	0.956035	0.969359	1.028914	1.034021	0.943013
	P8	1.040412	1.009744	1.028521	0.853652	1.330517
	P9	1.024026	1.036979	1.000857	1.100896	1.146319
	P10	1.008471	0.992312	0.966413	1.052351	0.614035
	P11	0.964083	0.977453	1.093296	0.96043	0.916661
	P12	0.959333	0.994604	0.945133	0.982212	1.011907
	P13	1.008457	1.049807	0.972897	1.05379	0.752243
	P14	1.031424	0.945	1.011352	0.943626	1.178505
	P15	1.018303	1.06908	0.963922	1.042131	1.354655
	P16	0.971546	0.986292	1.012766	1.038269	0.74655
	P17	0.918872	0.974066	1.13843	1.078726	0.695267
	P19	0.963535	1.013091	0.94508	0.961043	1.016437
	P20	1.099113	1.06623	0.966691	0.866844	0.964004
	P21	1.016019	1.060969	1.150757	0.932749	1.235064
	P22	0.92928	0.837175	0.759292	1.090259	0.733781
	P23	1.047871	1.113114	1.041855	1.04715	1.109723
	P24	1.002421	1.017206	1.040783	1.126387	1.07515
	P25	0.980157	0.945309	1.017099	0.943142	0.922725
	P26	1.03385	1.064558	0.900443	0.810683	0.986327
	P27	0.982795	1.003745	1.058596	1.015436	0.848095
	P28	1.024664	1.036086	0.978374	1.141728	1.192354
	P29	0.982984	0.991652	1.195466	0.942356	1.301605
	P30	1.009501	0.953863	1.200836	0.900781	0.991364
	P31	0.971978	1.024253	0.358829	1.031613	0.348933
	P32	1.02016	0.989398	1.214037	1.050052	1.079227
	P33	1.023821	1.008482	0.946818	1.038079	1.0067
	P34	1.025515	1.034438	1.11197	0.945225	1.413541
	P35	0.916834	0.957087	1.008396	1.043844	0.714557
	P36	0.992345	0.992586	0.92859	0.943933	0.958482
	P37	1.059354	1.120454	1.07858	1.004958	1.29519
	P38	0.939546	0.870139	0.928491	1.088188	0.696699
	P39	1.029746	0.986365	0.983728	0.981504	0.997079
	P40	1.006548	1.044174	1.021569	0.880405	1.107544
	P41	0.966024	1.072006	1.071261	1.08809	1.032669
	P42	1.038607	1.035276	1.026511	0.947132	1.113317
X-linked	ZIC3	0.903039	0.540452	0.571464	0.877582	0.656394
	P43	1.044737	1.088605	1.013242	0.882541	1.121619
	P44	1.01481	1.016545	1.052305	0.871781	0.924239
	P45	0.948352	0.876088	1.083451	1.213716	0.98852
	P46	0.985578	1.047308	0.847114	1.014807	0.862489
Y-linked	SRY	0.001588	1.21337	1.103978	0	1.193536
	IRX3	1.018346	1.148461	0.771363	0.965606	1.177464
	new18	1.017899	0.966462	1.205473	0.875317	1.096924
	P47	1.029168	1.027415	0.994514	0.959788	0.74729
	MEF2C	1.016275	0.978269	1.043895	1.08308	1.273262
	HEY2	0.876592	0.910877	1.003417	1.070803	1.382976
	P48	1.026109	1.021218	0.91956	0.936212	0.762253
	P49	0.952722	0.978715	1.145803	1.300348	1.236581

Figure 4.9: Representative example of data obtained using the Maphematica Excel macro. Columns 1 and 4 show a profile compatible with a female sample (signal from Y-linked probe close to zero and X-linked probe signal close to 1). Columns 2, 3 and 5 displayed male profile (an X-linked probe signal close to 0.5 and a Y-linked probe signal close to 1). Column 3 and column 4 additionally showed a signal reduction from probe 31 and a signal increase from probe 49 respectively compatible with copy number variation of the segment spanned by those probes that was not observed in subsequent analysis of those samples. Column 5 shows several numbers outside the reference interval. Detailed analysis revealed low quality data for low overall intensity of the relevant trace.

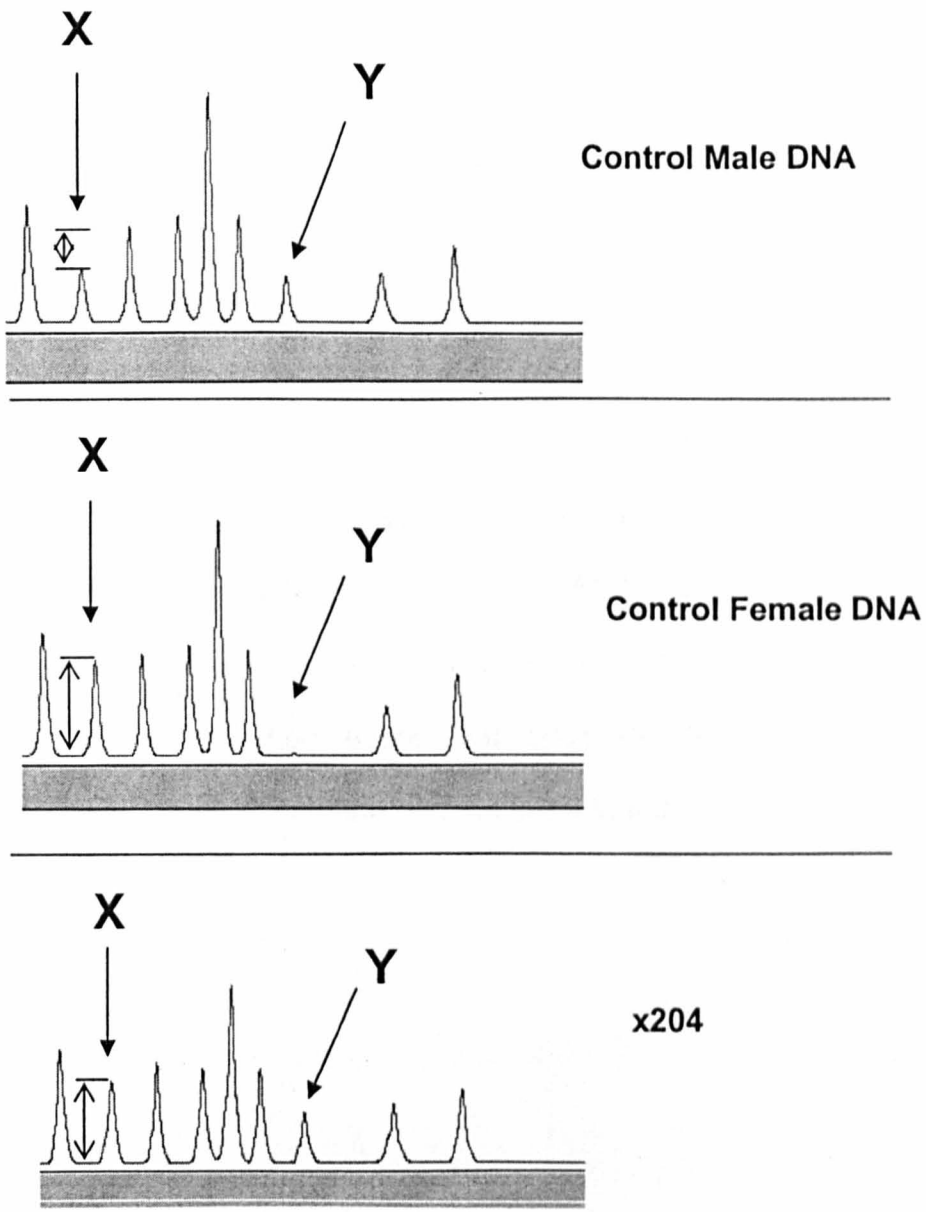


Figure 4.10: Partial MAPH trace of sample x204 sample showing an X-linked probe profile suggestive of disomy of chromosome X and normal male dosage of the Y-linked probe (XXY). This result was confirmed by repetition of the analysis with the same probe set and the XpYp probe set.

4.5 Discussion

The MAPH method was used in our congenital heart disease cohort to detect copy number variations of the *MYH6* gene. Partial duplications of this gene have been reported, in the form of a *MYH6/MYH7* hybrid gene flanked by normal copies of these two genes in cardiomyopathy patients. The complementary rearrangement, a *MYH7/MYH6* hybrid without normal flanking cardiac myosin gene is an obligate product of the non-allelic homologous recombination event that originated the cardiomyopathy hybrid gene, but such a mutation has not been found.

Probe	Type	Location	1	2	3	4	5	6	7	8	9	10	11	12	x204	F x204	M x204	Standard deviation
ST16A6	PA	XpYp	1.140341	1.059794	1.014051	1.000744	1.000473	1.002396	0.999898	1.022699	1.097412	1.095732	0.720733	1.030108	1.01649	1.019298	0.779831	0.110408083
ST5G1	A	15q	0.802519	0.810361	0.814286	0.91287	1.12173	0.947598	0.854687	0.851812	0.887626	0.934748	1.517559	0.925985	1.07258	0.945224	1.600415	0.244464511
ST16A1	PA	XpYp	1.108616	1.207755	0.982273	1.096475	0.834219	1.087542	1.135689	0.957224	1.090335	1.134791	0.761724	1.019628	0.87239	1.014545	0.696793	0.149257637
ST15C10	A	21q	0.901088	0.837356	0.994564	0.939146	0.989563	0.923415	0.950282	0.959175	0.954405	0.811459	1.324261	1.002187	1.126115	0.977876	1.309109	0.147551807
ST16A2	PA	XpYp	1.173231	1.177638	1.186416	1.03774	1.006456	1.009013	1.053319	1.154521	1.054089	1.052437	0.60515	1.005663	0.861664	0.999385	0.623277	0.178736177
ST18G1	A	12q	0.780209	0.747126	0.873196	0.907977	1.112576	0.953401	0.882102	0.977758	0.846853	0.912014	1.421652	0.974912	1.15189	1.021514	1.436819	0.205287262
2d2	X	Xq28	0.777323	0.830909	0.420124	0.905782	0.41805	0.99891	0.436749	0.450183	0.949009	0.473246	1.269845	0.847357	0.986257	0.440965	1.295291	0.392575607
a12	Y	SRY	0.134445	0.123272	1.002379	0.131573	0.927159	0.135431	1.003886	1.078585	0.089627	0.884694	0.152284	0.130452	1.083419	1.019878	0.15551	0.65836992
ST16A4	PA	XpYp	1.159138	1.124844	1.229487	1.149132	0.948738	0.957951	1.039255	1.009921	1.18102	1.159346	0.634405	0.952399	0.880477	1.040881	0.533006	0.198201969
PMP3	A	17	0.879426	0.828062	1.04588	1.005539	1.126577	0.908866	0.91453	1.028817	0.997785	0.959109	1.199963	0.933196	1.001817	1.062461	1.107973	0.100138013
ST17H1C	PA	XpYp	1.223382	1.297239	0.775565	0.802724	0.806848	1.181394	1.197026	1.054637	0.846866	0.895933	0.915656	1.16831	0.907623	0.850922	1.075877	0.178444187
ST17A1	A	1q	0.859419	0.932819	1.05029	1.185053	1.1156	1.061816	0.921973	0.838164	1.051975	1.011792	1.017129	0.940613	0.839433	1.050929	1.122996	0.107130718
PGPL2	PA	XpYp	0.848552	0.835719	1.060529	1.015615	1.118702	0.852589	0.880246	0.96974	1.048286	1.06696	1.071085	0.931043	1.333091	1.032585	0.935258	0.130311526
ST17G1	A	16p	0.98476	0.875088	1.044578	1.016982	0.958061	0.883059	0.870584	1.140685	0.991575	1.02302	1.075403	0.982783	1.082823	1.021159	1.049441	0.078567773
ST10D1	A	17q	0.938139	0.931655	1.173919	0.968034	1.074012	0.969155	1.133398	0.923073	1.104966	0.916538	1.037956	0.913465	0.927235	1.072743	0.915714	0.09044041
DMD53A	X	Xp21	0.74791	0.800897	0.561138	0.998783	0.538447	0.924306	0.474406	0.467352	1.006588	0.467379	1.082013	1.006596	0.979214	0.53816	0.906811	0.336730735
ST19A1	PA	XpYp	1.186432	1.167647	0.772423	1.136275	0.76098	1.203577	1.010316	0.972355	0.959337	1.213337	0.891312	0.994924	0.76001	0.946556	1.02452	0.158728694
ST18C1	A	9p	0.767169	0.824871	1.279374	1.07009	1.146972	0.78934	0.93239	0.945415	0.989771	0.978395	1.181664	0.999617	1.006819	1.170079	0.918033	0.148957712
ST19C1	PA	XpYp	1.274282	1.313968	0.703632	0.612484	0.903881	1.209863	1.20373	1.259635	0.739223	0.707985	0.716025	1.191456	1.279646	0.638964	1.245225	0.281991006
ZIC3	X	Xq25	0.823412	0.865125	0.523452	1.00162	0.535768	0.934739	0.472501	0.472144	1.129955	0.585934	0.964249	0.84469	0.932244	0.568906	0.845261	0.325274379
ST17C9	A	6q	0.886424	0.81371	1.077526	1.088482	1.093023	0.938103	0.945993	0.963757	0.993764	0.997037	1.049141	1.006415	1.085	1.086311	0.975315	0.082784295
ST18F2	A	7q	1.005905	1.033287	1.090711	1.12177	0.949886	0.972441	0.915172	0.863764	1.179377	1.061535	1.118247	0.91397	0.793277	1.084869	0.895789	0.110637681
ST4G4	A	3q	0.905625	0.887201	1.114496	0.957969	0.918569	0.936542	0.977138	1.04392	0.887885	0.915236	1.196847	1.00145	1.106436	1.045907	1.10478	0.097297727
ST17E3	A	10q	0.92245	0.976449	0.894924	0.969044	1.152343	0.975032	1.042216	0.96216	1.042097	1.03985	0.967997	0.973454	1.040649	0.942735	1.098599	0.068325092
ST16A8	PA	XpYp	1.225656	1.19905	0.980657	1.091301	0.994694	1.005056	1.039991	1.042509	1.097038	1.185047	0.566208	1.039185	0.913428	1.080174	0.540006	0.200285164
ST14A2	A	4p	0.953972	0.898369	0.966804	0.989176	1.063385	1.047073	0.947468	0.987199	0.952685	0.911579	1.085317	1.007972	1.115752	0.930102	1.143147	0.075118123
			FC	FC	MC	FC	MC	FC	MC	MC	FC	MC	FC	FC	MC	MC	FC	

Figure 4.11: Results of MAPH analysis of sample from 12 control subjects, patient x204 and his parents showing a normal male Y linked signal, a dosage of all three X-linked probes (Xp21, Xq26 and Xq28) and pseudoautosomal region 1(PAR1) probe PGPL2, the most centromeric of all PAR1 probes, compatible with a posible partial disomy of chromosome X in the male patient, excluding most of PAR1. PA, pseudoautosomal; A, autosomal; FC, female control; MC, male control.

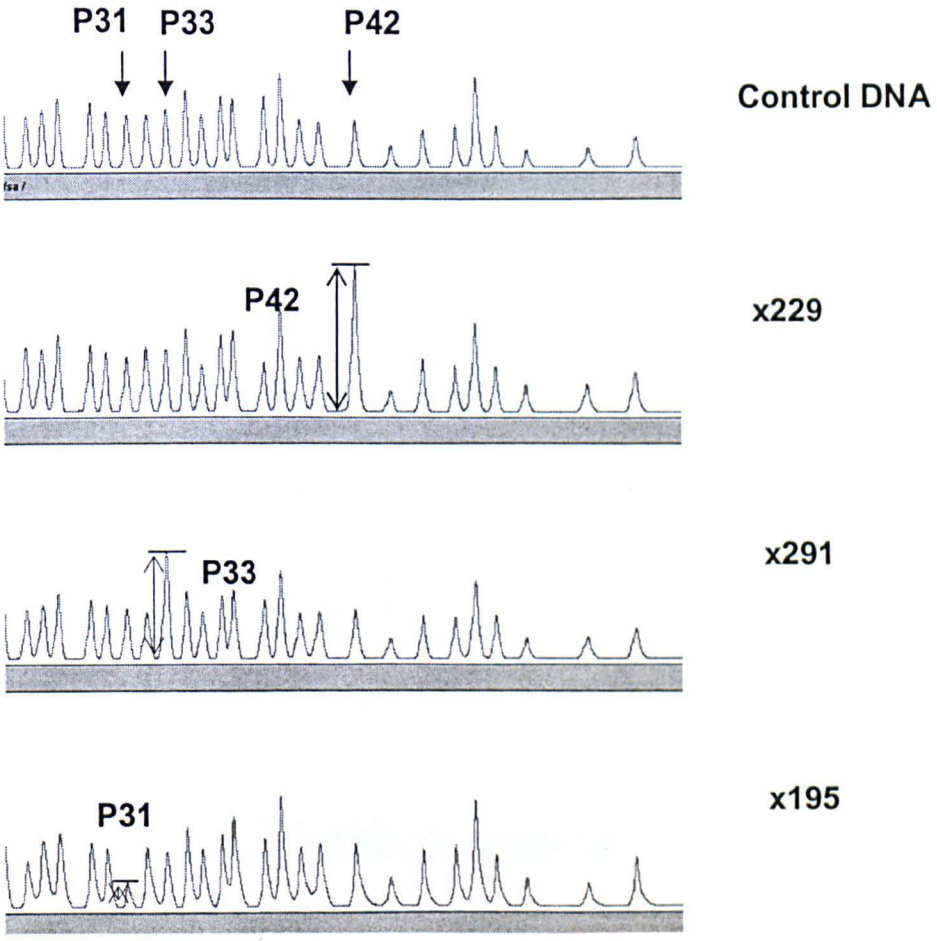


Figure 4.12: Representative examples of three MAPH traces suggestive of copy number variation. Samples x229 and x291 showed a relative increment and sample x195 a relative decrement of the peaks areas from probes P42, P33 and P31 respectively in comparison with control DNA traces. None of these variations were confirmed in subsequent MAPH assays from the same samples (see Figure 4.13).

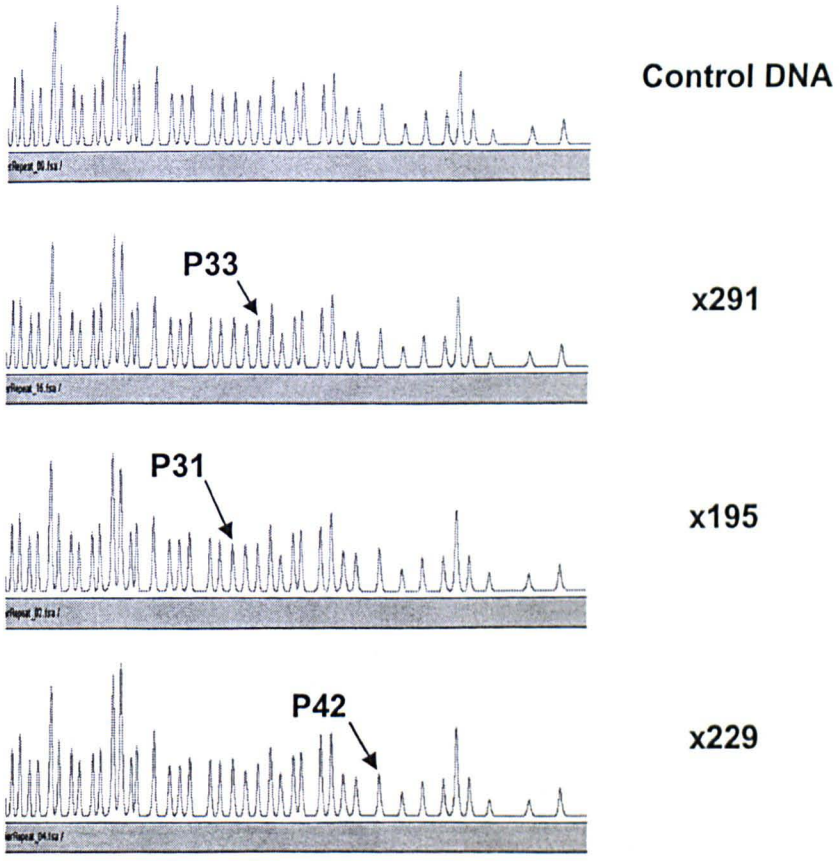


Figure 4.13: Subsequent MAPH assays of samples x291, x195 and x229. In the first round of MAPH analysis these samples showed traces compatible with copy number variation. As the abnormality in the traces failed to appear in two further repetitions of the whole process and in new secondary PCR and electrophoresis of the first round recovered probe, it was concluded that the first result was an electrophoresis artefact.

Although the absence of a MAPH result suggestive of deletion or duplication in our patients samples indicate that copy number variation of the *MYH6* gene is not a main contributor to sporadic CHD, it is important to notice that the *MYH6* MAPH probe set has some limitations. The high similarity between several myosin genes across the genome makes the task of designing a probe with no more than one potential cross-hybridization target very difficult in very well conserved segments.

MYH6 exons showing high similarity with other exons, in other myosin genes were not assayed directly. Instead, when possible, probes were designed in neighbouring introns (consider exon 15 or 26 of *MYH6* in Figure 4.2). If a deletion occurs within any of those exons it can not be detected with the *MYH6* probe set.

According to the advice of experienced MAPH users, false positive traces like those showed in Figure 4.12, are mostly electrophoresis artefacts and tend to occur in samples located in the first columns of a 96-well tray. Two of the three traces in the Figure were located in the first two columns of the tray during electrophoresis.

4.5.1 Detection of a case XXY

The fact that through the use of the *MYH6* probe set was possible to detect an XXY patient, indicates that the *MYH6* MAPH can allow us to distinguish real copy number variation. This variation was seen in three different MAPH experiments and verified using an entirely different probe set, XpYp, which contains multiple X- , Y-linked and pseudoautosomal (PAR1 only) probes. As expected, the signals of the Y- and X-linked probes were compatible with an XXY diagnosis; however, most of the pseudoautosomal probes revealed a normal profile. This suggests that the disomy of X in this patient is not complete and that a fraction (if not all) of the genes of the PAR1 region exist in normal dosage. If the genes escaping X inactivation are in normal dosage, the risk of the child of developing Klinefelter stigmata could be considerably reduced.

Chapter 5 THE SEARCH FOR A NEW HOLT-ORAM GENE

5.1 Introduction

A high proportion of Mendelian syndromes affecting the limbs include as a feature a cardiac defect (Table 5.1).

The term heart-hand syndrome is reserved to entities where the main features occur in heat and upper limb with relatively minor involvement of other structures.

Holt-Oram syndrome is the more common of the heart-hand syndromes. It shows an autosomal dominant pattern of inheritance and it is observed in 1 on 100,000 live births (Elek et al. 1991). This syndrome was first recognized in a 4 generation family with atrial septal defect, arrhythmia, various ECG alterations and variable degrees of upper limb abnormality, ranging from absent and triphalangeal thumb to radial aplasia and clavicle malformation (Holt and Oram 1960). Multiple familial cases were reported soon after, showing great phenotypic variability (Holmes 1965; Lewis et al. 1965b; Zeterqvist 1963).

Table 5.1: Mendelian syndromes with mayor cardiac and limb features.

GENE	PHENOTYPE	CARDIAC DEFECT
Common		
unknown	Adams-Oliver Syndrome	ASD, double-outlet right ventricle, TOF, VSD, PS, aortic coarctation, aortic valvar stenosis, mitral valvar prolapse
unknown	Goltz Syndrome	PS, ASD,PAD
TBX5	Holt-Oram Syndrome	ASD,VSD, TOF, AS
Unknown	Poland Syndrome	Dextrocardia, stenosis of left subclavian artery
Unknown	Thrombocytopenia-absent radius Syndrome	TOF, ASD
Unknown	VACTERL Syndrome	VSD, TOF, PAD, transposition of the great arteries
Unknown	Carpenter Syndrome	VSD, PAD, ASD,PS, TOF
Unknown	Heart-hand IV Syndrome	PS, VSD, common atrium, anomalous systemic venous drainage
Unknown	Hydroletalus Syndrome	AVSD, VSD
MKS1	Meckel Syndrome	Septal defects, aortic coarctation, PAD
Rare		
Unknown	Brachydactyly type E with ASD Syndrome	Secundum type ASD
Unknown	Hajdu-Cheney Syndrome	PAD, VSD, AS, mitral incompetence
Unknown	Hollister Syndrome	PS, conduction defects
Unknown	Pena-Shokeir Syndrome	ASD, aortic coartation, hypoplastic heart
Unknown	Ter Haar Syndrome	Double outlet right ventricle, PAD, VSD, mitral valvar prolapse
Unknown	Ho Syndrome	VSD, aberrant subclavian artery
Unknown	Holzgreve Syndrome	ASD, VSD, hypoplastic left heart, atretic aortic arch
Unknown	Jeune thoracic dysplasia	Pulmonary hypoplasia
BBS6	Kaufman-McKusick Syndrome	ASD, VSD, single atrium
Unknown	Laurence Syndrome	ASD, VSD, Mitral stenosis, aortic coarction
Unknown	Oral-cardiac-digital Syndrome	AVSD, coarction/interrupted aorta, aortic incompetence
GL13	Pallister-Hall Syndrome	PS, ASD, VSD, aortic coarction
Unknown	Schinzel-Giedion Syndrome	ASD, PAD, valvar anomalies
GPC3	Simpson-Golabi-Behmel Syndrome	VSD, ASD, PS, conduction defects
Unknown	Varadi Syndrome	ASV
Unknown	Young-Madders Syndrome	ASD, VSD, AVSD. Right-sided heart

TOF= tetralogy of Fallot, PS= pulmonary stenosis, PAD=persistent arterial duct, AS= aortic stenosis, AVSD=atrioventricular septal defect

5.1.1 Holt-Oram syndrome

Secundum atrial and ventricular septal defects are the most common cardiac abnormalities, followed by mitral valve prolapse, tetralogy of Fallot, aortic stenosis, dextrocardia, pulmonary stenosis, patent ductus arteriosus, amongst others. (Newbury-Ecob et al. 1996; Smith et al. 1979).

The skeletal abnormalities are the most constant feature of the disease, they can range from minor thumb defects to phocomelia and tend to be bilateral and more severe in the left upper limb. Thumb abnormalities are the most common defects. The thumb can be absent, hypoplastic or triphalangeal. Fingers are often short, and can sometimes be absent (especially the second and third fingers) if the thumb is also absent. The carpal bones can be absent fused or irregular or supernumerary bones may exist. Radial hypoplasia and aplasia are not uncommon and are accompanied by ulnar hypoplasia. The humerus and clavicles can be hypoplastic (Basson et al. 1994a; Newbury-Ecob et al. 1996; Poznanski et al. 1970; Smith et al. 1979) (Figures 5.1 and 5.2).

Figure 5.1: Limb abnormalities in a Holt-Oram syndrome patient. Hypoplastic radius and ulna, absence of thumb and first metacarpal bone. Figure taken from emedecine.com.

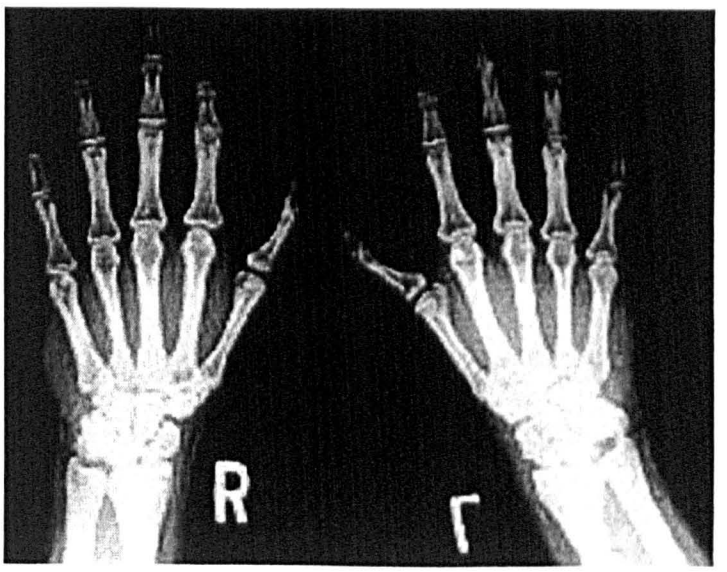


Figure 5.2: Radiograph of the hands of a patient with Holt-Oram syndrome. The distal phalanx of the left thumb is hypoplastic. Carpal abnormalities are more pronounced in left side and consist of deformity with enlargement of scaphoid and trapezium. Source: Craig T Basson, Deborah A McDermott. Cardiovascular Research, Greenberg Division of Cardiology, Department of Medicine, Weill Medical College of Cornell University.

5.1.2 Identification of *TBX5* with HOS1

Linkage studies were conducted in large Holt-Oram families by several groups in order to identify the genomic location of the gene responsible for Holt-Oram syndrome (*HOS1*) (Basson et al. 1994b; Bonnet et al. 1994; Terrett et al. 1994). They located a critical interval in the long arm of chromosome 12q21.3-12q22.

An affected patient carrying a complex rearrangement of chromosome 12 was identified (Terrett et al. 1996), with two of the breakpoints within the critical interval defined by linkage. This allowed the gene identification efforts to focus in a shorter segment.

By means of exon trapping, database searching of EST sequences and cDNA library screening, two groups independently located the human orthologues of mouse Tbx5 and Tbx3 and identified mutations by SSCP in the human TBX5 in Holt-Oram patients (Basson et al. 1997; Li et al. 1997).

5.1.3 Mutations of TBX5

Several additional mutations (Table 5.2) have been reported by other groups (Akrami et al. 2001; Basson et al. 1999; Brassington et al. 2003;

Cross et al. 2000a; Fan et al. 2003; Gruenauer-Kloevekorn and Froster 2003; Heinritz et al. 2005; Reamon-Buettner and Borlak 2004; Yang et al. 2000).

It was observed that in a family with a mutation in the N-terminal end of the T-Box domain of TBX5 (G80R) the cardiac abnormalities predominate over the limb defects and that murine Tbx5 with this mutation failed to interact with Nkx2.5 and activate the cardiac specific *Nppa* gene in COS-7 transfection experiments.

In subjects with the R237Q and R237W mutations (C-terminal end of the T-box) it was noticed that severe limb defects coexisted with relatively mild cardiac abnormalities. The mutant Tbx5 with the former mutation was capable to activate Nppa in a level comparable to the wild-type protein in the same transfection experiments (Basson et al. 1999; Hiroi et al. 2001).

These observations suggested a genotype-phenotype correlation; however, a study taking in to account fourteen different mutations in 17 families failed to confirm a clear tendency (Brassington et al. 2003).

Table 5.2: Mutations of TBX5

Mutation	Type	Predicted amino acid change	Reference
100delG	FS	A34P→65X	7
100-101insG	FS	A34G→60X	7
145 C→A	MS	Q49K	5
161 T→C	MS	I54T	5
161 T→C (somatic)	MS	I54T	10
192 G→A	NS	W64X	8
205 G→T	NS	E69X	1
236 C→T (somatic)	MS	A79V	10
238 G→A	MS	G80R	3
246delGinsAA	FS	R82R→95X	3
280delC	FS	L94L→123X	11
287 C→T (somatic)	MS	P96L	10
299 A→G (somatic)	MS	P96L	10
305 T→C (somatic)	MS	L102P	10
361 T→G	MS	W121G	7
376-402del27bp	IFD	del146 (KAEPAMPGR)	8
400-401insC	FS	R134P→182X	7
400 C→T(somatic)	MS	R134C	10
408 C→A	NS	Y136X	9
416delC	FS	P139Q→149X	5
420-432del13bp	FS	D140E→145X	3
426-427insC	FS	p.A143fsX182	7
431 C→T(somatic)	MS	T144I	10
456delC	FS	L152L→173X	7
467-468insA	FS	Q156Q→182X	11
(439-484dup)	FS	ins162(HWMRQLVSFQKLKLT)→197X	11
505 G→A	MS	G169R	4
568 G→T	NS	E190X	4
584 G→C	MS	G195A	7
587 C→A	NS	S196X	2
593-594insA	FS	N198K→208X	2
668 C→T	MS	T223M	7
709 C→T	MS	R237W	3
710 G→A	MS	R237Q	1
727delG	FS	E243S→263X	3
755 G→T	MS	S252I	4
781 A→T	MS	S261C	7
797 A→G(somatic)	MS	K266R	10
798delA	FS	K266K→393X	7
805delT	FS	S269P→393X	4
835 C→T	NS	R279X	2
875 A→G(somatic)	MS	Q292R	10
946 G→T	NS	E316X	4
1084 C→T	NS	Q362X	11
1159-1160insA	FS	S387K→486X	2
del Exons 3-9	LD	del(Exon3-9)	6

In the first column the numbers indicate the position in sequence Accession number U8953. In the third column, for frameshift deletions the element to the left of the arrow refers to the first modified residue; the element to the right refers to the amino acid position were the aberrant stop codon occurs. Numbers indicate the position of the change in sequence Accession number AAC04619. Ins, insertion; del, deletion; dup, duplication; FS, frameshift; MS, missense; NS, nonsense; IFD, in-frame deletion; LD, large deletion. References: 1)Basson, Bachinsky et al. 1997; 2) Li, Newbury-Ecob et al. 1997; 3) Basson, Huang et al. 1999; Cross, Ching et al. 2000; 5) Yang, Hu et al. 2000; 6) Akrami, Winter et al. 2001; 7) Brassington, Sung et al. 2003; 8) Fan, Duhagon et al. 2003; 9) Gruenauer-Kloevekorn and Froster 2003; 10) Reamon-Buettner and Borlak 2004; 11) Heinritz, Moschik et al. 2005

Traditionally, the sensitivity of the molecular testing for *TBX5* mutations in Holt-Oram patients ranged from 22% to 35% (Brassington et al. 2003; Cross et al. 2000b; Heinritz et al. 2005). However, when only patients complying with the minimal criteria (Newbury-Ecob et al. 1996), i.e. at least one family member with radial ray defect and septal defects (atrial or ventricular) or atrioventricular block, are screened, the sensitivity of the test exceeds 70% (McDermott et al. 2005; Mori and Bruneau 2004).

5.1.4 Other loci involved

The finding of a Holt-Oram patient with a pericentric inversion with a breakpoint at chromosome 20q13 (Yang et al. 1990) led two groups to find, in Holt-Oram cases, mutations in *SALL4*, a gene located in that region and previously known to cause Okihiro syndrome, a phenotype that can overlap with Holt-Oram syndrome (Brassington et al. 2003; Kohlhase et al. 2003).

Two deletions of chromosome 14q (Le Meur et al. 2005; Turleau et al. 1984), a large deletion of chromosome 4q (Ockey et al. 1967), and

chromosome 5 (Rybak et al. 1971) have been reported in patients with heart-hand syndromes resembling Holt-Oram syndrome in variable degrees, but no mutations have been found in those intervals.

5.1.5 Previous analysis of chromosome 12q non-linked families

When the initial linkage analysis to locate HOS1 (*TBX5*) was carried out, two Holt-Oram, families (HOS15 and HOS32, Figures 5.3 and 5.4) that did not show linkage to markers in chromosome 12q, were identified (Terrett et al. 1994). DNA samples from individuals of family HOS15 were used to perform a genome-wide linkage analysis, in order to locate a second gene responsible for Holt-Oram syndrome (HOS2). After an initial, low resolution analysis of 400 microsatellites, fine mapping was carried out in five candidate regions (Figure 5.5) that showed the highest LOD scores in chromosomes 4, 8, 11, 16 and 17 (Ching 2001; Cross 2003).

The analysis of the haplotypes obtained by genotyping the microsatellites within the candidate intervals revealed that for the regions in chromosomes 8, 11, 16 and 17 more than one unaffected individual in the family shared the alleles present in all the affected subjects and therefore they needed to be considered as "non-penetrant" cases if each interval is considered to include HOS2. The same happened for just one individual for the region in chromosome 4 (Figure 5.3). The mapping efforts were focused in the 4q interval, that showed the highest LOD score 1.75. A

minimum (from marker D4S413 to D4S2962) and maximum (D4S2962 to D4S1595) linkage intervals were defined (Cross 2003).

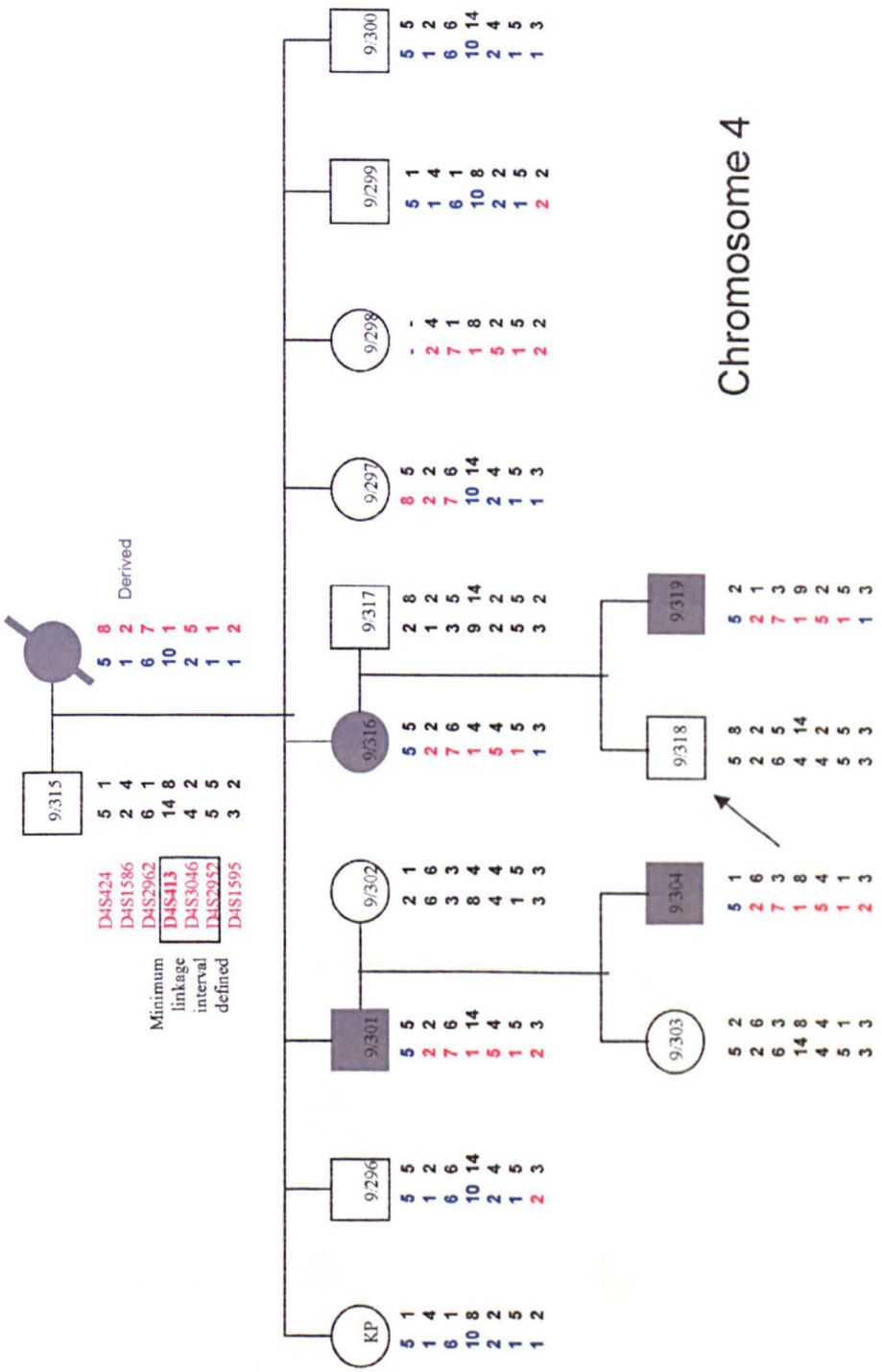


Figure 5.3: HOS15 family. Haplotypes of markers in the HOS2 candidate region 4q. The haplotypes of the affected founder are derived. The "affected haplotype" is shown in red. The arrow points to subject 9/318. Modified from Cross 2003.

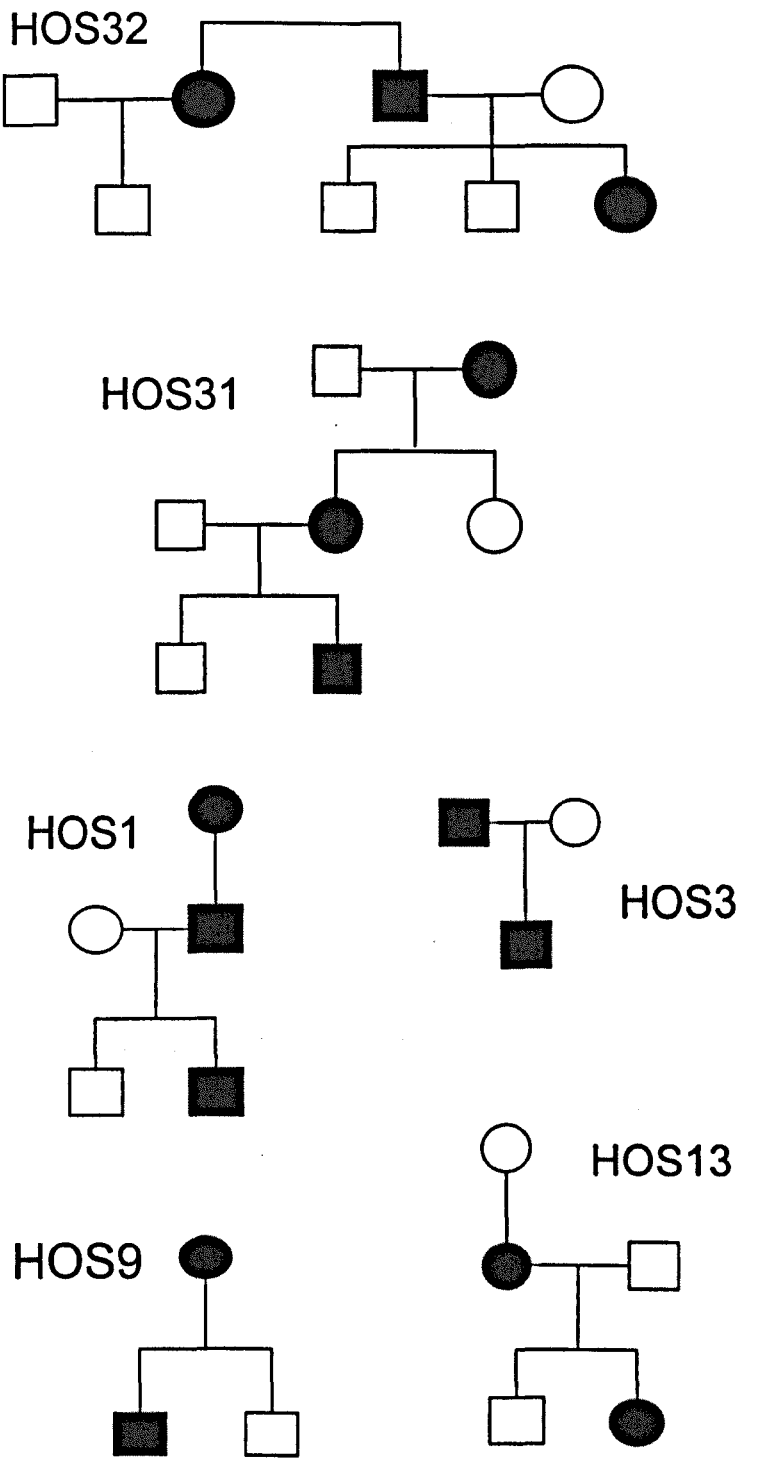


Figure 5.4: Pedigrees of other HOS families not linked to chromosome 12q.

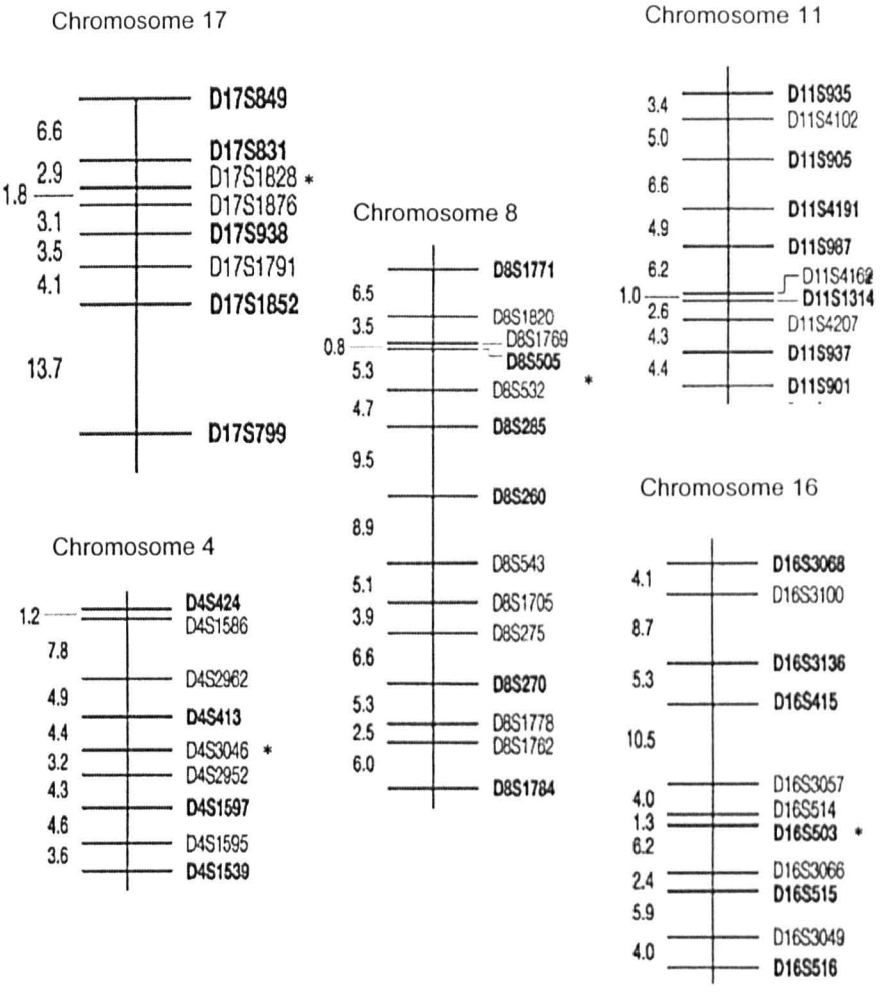


Figure 5.5: Microsatellite markers used by S. Cross for fine mapping of regions with the highest LOD sore obtained during his genome-wide linkage analysis. Taken from Cross 2003.

Analysis by SSCP was carried out in three genes within or close the 4q interval, *MADH1*, *TLL1* and *HAND2*, but no mutation was found (Cross 2003).

5.2 Aim of the project

The aims of the project described in this chapter were: a) to reduce the maximum linkage interval in the 4q candidate region, b) to systematically screen the expression pattern of the genes in the new maximum interval to prioritize targets for mutational analysis and c) to find a new Holt-Oram syndrome causing gene.

5.3 Methods

5.3.1 Physical annotation of the linkage intervals

The sequence of the microsatellites used to define the maximum and minimum linkage intervals were retrieved from GeneBank at the NCBI website. These sequences were used as BLAST queries for comparisons against the sequenced human genome (Build 36).The chromosomal

localization of the BLAST hits and surrounding genomic features was established using the Genome View interface of the BLAST output.

Microsatellite D4S2952 defines the centromeric limit of the maximum linkage interval and its sequence (Accession: Z53372) maps approximately to 150,580kb of chromosome 4 (pter-qter) and is located within the second last intron of the LOC285423, a genomic alignment match to IMAGE clone 5295442.

D4S413 (Accession: Z16837) defines the centromeric end of the minimum interval and maps to 158,570kb. It is located between the ionotropic glutamate receptor gene (*GRIA2*) and LOC729937, a computational prediction of a transcript by the program GNOMON that resembles and is possibly related to the *MAST2* gene (microtubule associated serine/threonine kinase 2) located in 1p34.1.

D4S2952 (Accession: Z51120), at the telomeric end of the minimum interval is located between the carboxipeptidase E gene (*CPE*) and a GNOMON prediction (LOC402191), possibly a pseudogene similar to nucleolar protein 8. It is approximately located at 166,675kb.

At the telomeric limit of the maximum interval, D4S1595 (Accession: Z24146) is located in the second intron of *SCRG1* (scrapie responsive protein 1) and maps to approximately 174,554kb from the pter end of chromosome 4.

Knowing the physical location of these microsatellites, it was possible to calculate the length of the minimum (8.105Mb) and maximum (23.974Mb) linkage intervals. A list of "gene-related" genomic features in the whole maximum interval was made using the Map Viewer interface at the NCBI website (Table 5.3), incorporating known genes, alignments of well characterized cDNA clones, mRNAs and ESTs as well as GNOMON predictions. A total of 126 gene related elements were located within the maximum interval and 43 within the minimum interval.

Additional microsatellites were located between the limits of the minimum and the maximum interval in order to define a smaller maximum interval to reduce the number of candidate genes for expression and eventually mutational analysis.

By the time the analysis of the 4q region was stopped (Section 5.3.5.1), one microsatellite (D4S1646) had been typed (Figure 5.6).

Typing of the four microsatellites defining both intervals, plus those to be used for reduction of the maximum interval, was planned on subjects belonging to families presumably not linked to chromosome 12q but too small to be used in a genome-wide analysis. The purpose was to test them for consistency with a possible linkage with chromosome 4q. By the time the analysis of the 4q region was abandoned (see Section 5.3.5.1),

Table 5.3 continues in next page....

Symbol	Description
D4S2962	**Microsatellite: centromeric end of maximum linkage interval**
LOC285423	hypothetical gene supported by BC031092
DCAMKL2	doublecortin and CaM kinase-like 2
LRBA	LPS-responsive vesicle trafficking, beach and anchor containing
LOC729558	hypothetical protein LOC729558
MAB21L2	mab-21-like 2 (C. elegans)
LOC729566	similar to zinc finger and BTB domain containing 8 opposite strand
LOC649288	similar to Adenylate kinase isoenzyme 4, mitochondrial (ATP-AMP transphosphorylase)
RPS3A	ribosomal protein S3A
SNORD73B	small nucleolar RNA, C/D box U73B pseudogene
SNORD73A	small nucleolar RNA, C/D box 73A
SH3D19	SH3 domain protein D19
ESSPL	epidermis-specific serine protease-like protein
LOC729830	similar to CG3558-PA, isoform A
PET112L	PET112-like (yeast)
FBXW7	F-box and WD-40 domain protein 7 (archipelago homolog, Drosophila)
DKFZP434I0714	hypothetical protein DKFZP434I0714
LOC391706	similar to 40S ribosomal protein S3a (V-fos transformation effector protein)
LOC646737	similar to ribosomal protein S14
TMEM154	transmembrane protein 154
TIGD4	tigger transposable element derived 4
ARFIP1	ADP-ribosylation factor interacting protein 1 (arfaptin 1)
LOC152667	NIP30-like
LOC729870	hypothetical protein LOC729870
KIAA1727	KIAA1727 protein
TRIM2	tripartite motif-containing 2
ANXA2P1	annexin A2 pseudogene 1
MND1	meiotic nuclear divisions 1 homolog (S. cerevisiae)
KIAA0922	KIAA0922
WDR45p	WDR45 psuedogene
TLR2	toll-like receptor 2
RNF175	ring finger protein 175
SFRP2	secreted frizzled-related protein 2
DCHS2	dachsous 2 (Drosophila)
PLRG1	pleiotropic regulator 1 (PRL1 homolog, Arabidopsis)
FGB	fibrinogen beta chain
FGA	fibrinogen alpha chain
FGG	fibrinogen gamma chain
LOC729651	similar to lecithin retinol acyltransferase (phosphatidylcholine--retinol O-acyltransferase)
LRAT	lecithin retinol acyltransferase (phosphatidylcholine--retinol O-acyltransferase)
MGC27016	hypothetical protein MGC27016
D4S2976	**Microsatellite**
LOC729902	hypothetical protein LOC729902
NPY2R	neuropeptide Y receptor Y2
MAP9	microtubule-associated protein 9
GUCY1A3	guanylate cyclase 1, soluble, alpha 3
GUCY1B3	guanylate cyclase 1, soluble, beta 3
ACCN5	amiloride-sensitive cation channel 5, intestinal
TDO2	tryptophan 2,3-dioxygenase
CTSO	cathepsin O
FTHP2	ferritin, heavy polypeptide pseudogene 2
LOC729923	similar to solute carrier family 22 (organic cation transporter), member 4
LOC646865	hypothetical protein LOC646865
PDGFC	platelet derived growth factor C
GLRB	glycine receptor, beta
LOC391707	similar to Chromatin accessibility complex protein 1 (CHRAC-1)
GRIA2	glutamate receptor, ionotropic, AMPA 2
D4S413	**Microsatellite: centromeric end of minimum linkage interval**
LOC729937	similar to microtubule associated serine/threonine kinase 2
C4orf18	chromosome 4 open reading frame 18
TMEM144	transmembrane protein 144
LOC646890	hypothetical protein LOC646890
RXFP1	relaxin/insulin-like family peptide receptor 1
LOC201725	hypothetical protein LOC201725
ETFDH	electron-transferring-flavoprotein dehydrogenase
PPID	peptidylprolyl isomerase D (cyclophilin D)
LOC729951	hypothetical protein LOC729951
FLJ25371	hypothetical protein FLJ25371

LOC338095	proteasome activator subunit 2 pseudogene
RAPGEF2	Rap guanine nucleotide exchange factor (GEF) 2
FSTL5	follistatin-like 5
LOC729725	hypothetical protein LOC729725
LOC729971	hypothetical protein LOC729971
LOC92345	hypothetical protein BC008207
LOC133332	mitochondrial ribosomal protein S5 pseudogene
NPY1R	neuropeptide Y receptor Y1
LOC729743	similar to neuropeptide Y receptor Y1
NPY5R	neuropeptide Y receptor Y5
TKTL2	transketolase-like 2
FLJ11184	hypothetical protein FLJ11184
MARCH1	membrane-associated ring finger (C3HC4) 1
LOC646954	similar to 14-3-3 protein theta (14-3-3 protein tau) (14-3-3 protein T-cell) (HS1 protein)
ANP32C	acidic (leucine-rich) nuclear phosphoprotein 32 family, member C
LOC653794	similar to ring finger protein 129
LOC646966	similar to 60S ribosomal protein L26
LOC391710	similar to ring finger protein 129
LOC389240	similar to nascent polypeptide-associated complex alpha polypeptide
LOC391711	hypothetical LOC391711
TRIM61	tripartite motif-containing 61
FLJ31659	hypothetical protein FLJ31659
LOC391713	similar to ring finger protein 129
LOC646989	similar to ring finger protein 129
TRIM60	tripartite motif-containing 60
TRIM75	tripartite motif-containing 75
FLJ38482	hypothetical protein FLJ38482
KLHL2	kelch-like 2, Mayven (Drosophila)
GKP3	glycerol kinase pseudogene 3
SC4MOL	sterol-C4-methyl oxidase-like
CPE	carboxypeptidase E
MIRN578	microRNA 578
HADHAP	(pseudogene) hydroxyacyl-Coenzyme A dehydrogenase
D4S2952	Microsatellite: telomeric end of minimum critical interval
LOC402191	similar to nucleolar protein 8
LOC646995	similar to Tripartite motif protein 38 (RING finger protein 15) (Zinc finger protein RoRet)
TLL1	tolloid-like 1
SPOCK3	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3
ANXA10	annexin A10
FLJ20035	hypothetical protein FLJ20035
FLJ31033	hypothetical protein FLJ31033
PALLD	palladin, cytoskeletal associated protein
LOC727835	similar to 60S ribosomal protein L9
CBR4	carbonyl reductase 4
SH3RF1	SH3 domain containing ring finger 1
NEK1	NIMA (never in mitosis gene a)-related kinase 1
CLCN3	chloride channel 3
C4orf27	chromosome 4 open reading frame 27
LOC441050	similar to unactive progesterone receptor, 23 kD
LOC402192	similar to Methylosome subunit pICln
MFAP3L	microfibrillar-associated protein 3-like
AADAT	aminoadipate aminotransferase
LOC642544	hypothetical protein LOC642544
LOC727891	similar to Heat shock protein HSP 90-alpha (HSP 86) (NY-REN-38 antigen)
D4S1646	Microsatellite
LOC441052	hypothetical gene supported by AF131741
GALNT17	polypeptide N-acetylgalactosaminyltransferase 17
FLJ46681	hypothetical protein LOC642588
GALNT7	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7
LOC727929	hypothetical protein LOC727929
HMGB2	high-mobility group box 2
SAP30	Sin3A-associated protein, 30kDa
SCRG1	scrapie responsive protein 1
D4S1595	Microsatellite: telomeric end of maximum linkage interval

Table 5.3: Gene-related features in the maximum and minimum linkage interval of chromosome 4q. The minimum interval is highlated in grey and the microsatellites in black.

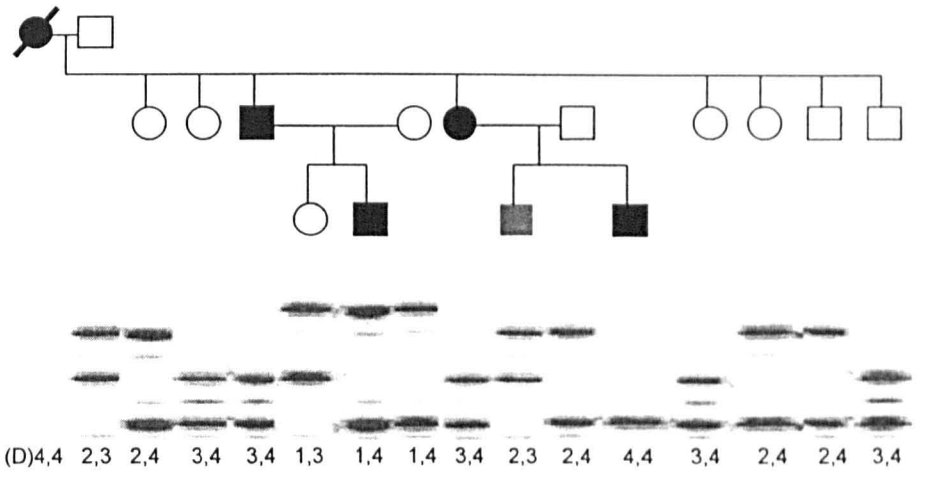


Figure 5.6: Fluorescent gel electrophoresis used for genotyping of the D4S1646 marker in subjects from family HOS15 in an attempt to reduce the length of the maximum linkage interval defined for chromosome 4q. Most of the meioses are uninformative for this marker as the deduced genotype of the affected founder (deceased) is 4,4.

typing of marker D4S2962 was done in those subjects and it is shown in Figure 5.7.

5.3.2 Identification of cDNA clones for expression analysis.

In order to assess if the expression pattern of the elements in the linkage interval resemble the expression pattern of *Tbx5* during embryonic development, BLAST comparisons were made between each of the gene-related features in the minimum interval against mouse ESTs using the known or predicted sequence of the relevant transcript as a query. The ESTs with the longest BLAST hit were selected. Expression analysis of the elements in the maximum interval would depend on its reduction by means of further microsatellite typing. Until the efforts in the 4q region were suspended (see Section 5.3.5.1), the idea was to use the mouse cDNA clones corresponding to the identified ESTs to make radioactive probes to be employed in Northern blot experiments.

The relevant IMAGE clones were requested to the MRC Geneservice (Cambridge, UK). On arrival, agar Petri dishes with the appropriate antibiotic were inoculated with material from the culture received, to obtain isolated colonies after overnight incubation at 37°C. Overnight broth antibiotic supplemented cultures were inoculated with material of the single colonies and located in an orbital incubator at 37°C, 200rpm. Minipreparations of plasmid DNA were made from the broth culture using the Qiagen Miniprep kit according to manufacturer's indications. Purified

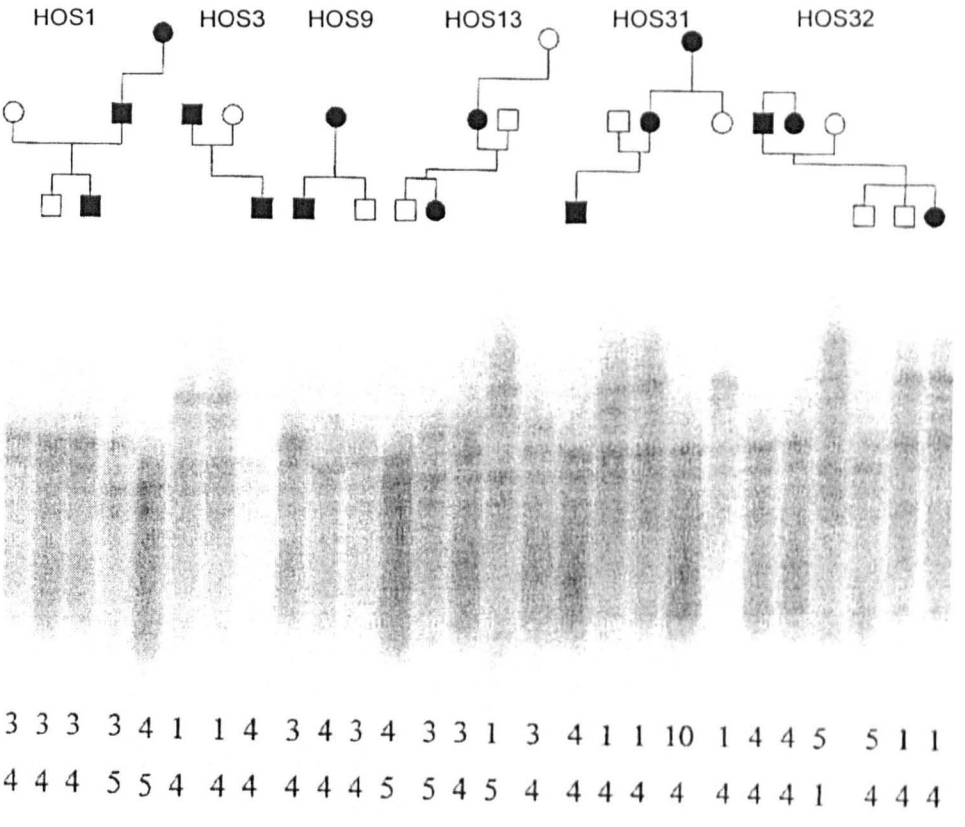


Figure 5.7: Radioactive polyacrilamyde gel electrophoresis used for genotyping of the D4S2962 marker in subjects from five families that showed no linkage with chromosome 12q in order to test for consistency with linkage to chromosome 4q. The allele calling is shown in the inferior two rows of numbers. Families HOS13 and HOS31are consistent with linkage between the disease and the marker. HOS1, HOS9 and HOS32 appear to be non-consistent with linkage and HOS3 is not informative for this marker.

plasmid DNA was used in double or single endonuclease restriction digestions with the appropriate enzymes (Table 5.4) in order to release the insert from the plasmid. A representative selection of agarose gel electrophoresis of the restriction digestion reactions is shown in Figure 5.8.

Table 5.4. cDNA clones for gene expression analysis

ELEMENT	IMAGE ID	DIGESTION WITH	
AD021	3026048	NotI	EcoRI
GRIA2	6516365	NotI	EcoRV
FLJ35882	1550743	NotI	EcoRI
LOC285505	3664790	XhoII	EcoRI
KLHL2	3987559	NotI	Sal I
LOC285504	6056308	NotI	Sal I
FLJ34659	752523	NotI	EcoRI
FLJ11155	3593194	SalI	NotI
LGR7	6051251	NotI	Sal I
LOC201725	6309627	NotI	EcoRV
ETFDH	6442431	NotI	Sal I
PPID	6507446	NotI	EcoRV
KIAA1450	555113	Sal I	NotI
LOC152940	6369005	XhoII	EcoRI
LOC152941	3025655	NotI	EcoRI
PDZ-GEF1	5699825	NotI	EcoRI
DKFZp566D234	6390044	Not I	EcoRV
LOC92345	6489898	Sal I	NotI
LOC133332	5146817	NotI	EcoRI
NPY1R	6490858	NotI	Sal I
NPY5R	461223	SfiI	SfiI
DKFZP434L1717	5165431	PCR used to produce insert DNA	
FLJ11184	5120919	NotI	EcoRI
FLJ20668	3812896	Sal I	NotI
LOC132689	2398278	NotI	EcoRI
LOC152682	4936600	NotI	Sal I
LOC166655	777988	Sal I	Mlu I
LOC201931	5328056	Sal I	NotI
SC4MOL	6336341	Not I	EcoRV
CPE	5707928	NotI	EcoRI

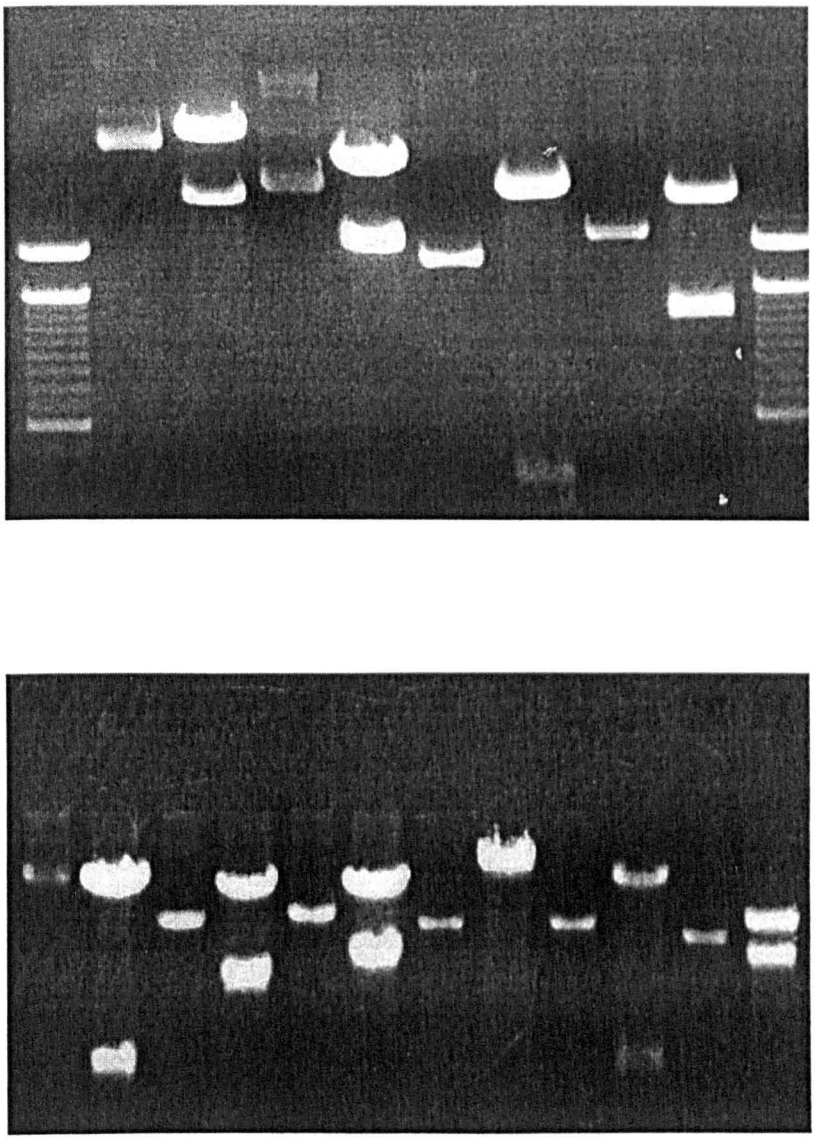


Figure 5.8: Representative agarose gel electrophoresis of IMAGE clones endonuclease restriction digestions performed to release the inserts. Each pair of contiguous lanes shows, to the left, undigested miniprep; to the right, a superior band of linearized vector and an inferior band, the released insert. Insert DNA was purified from excised agarose blocks containing the inferior bands.

5.3.3 RNA preparation

RNA to be used in the Northern blot experiments was prepared as follows: Total RNA was purified from forelimb, hindlimb and heart from 11.5dpc mouse embryos. The concentration and purity of the RNA was assessed in a Nanodrop spectrophotometer and agarose gel electrophoresis (Figure 5.9A). As an internal control for RNA quality, RT-PCR was carried out using oligo-dT for cDNA synthesis and mouse Tbx5 specific PCR primers (provided by Liz Packham). Amplification occurred as expected in RNA samples of all three tissues (Figure 5.9B).

5.3.4 Northern blot

The conditions and reagents to be used in the Northern blot were tested to ensure a rational use of the RNA extracted from mouse embryos. For that purpose, a riboprobe (provided by Thelma Robinson) produced using T7 RNA polymerase and a cDNA Tbx5 clone with the T7 phage promoter, was loaded in different concentrations in an agarose gel. The RNA was transferred and fixed to a nylon membrane and a Northern hybridization using probes generated with the same cDNA clone by the random priming method was carried out (Figure 5.10).

5.3.5 Study in 16q region

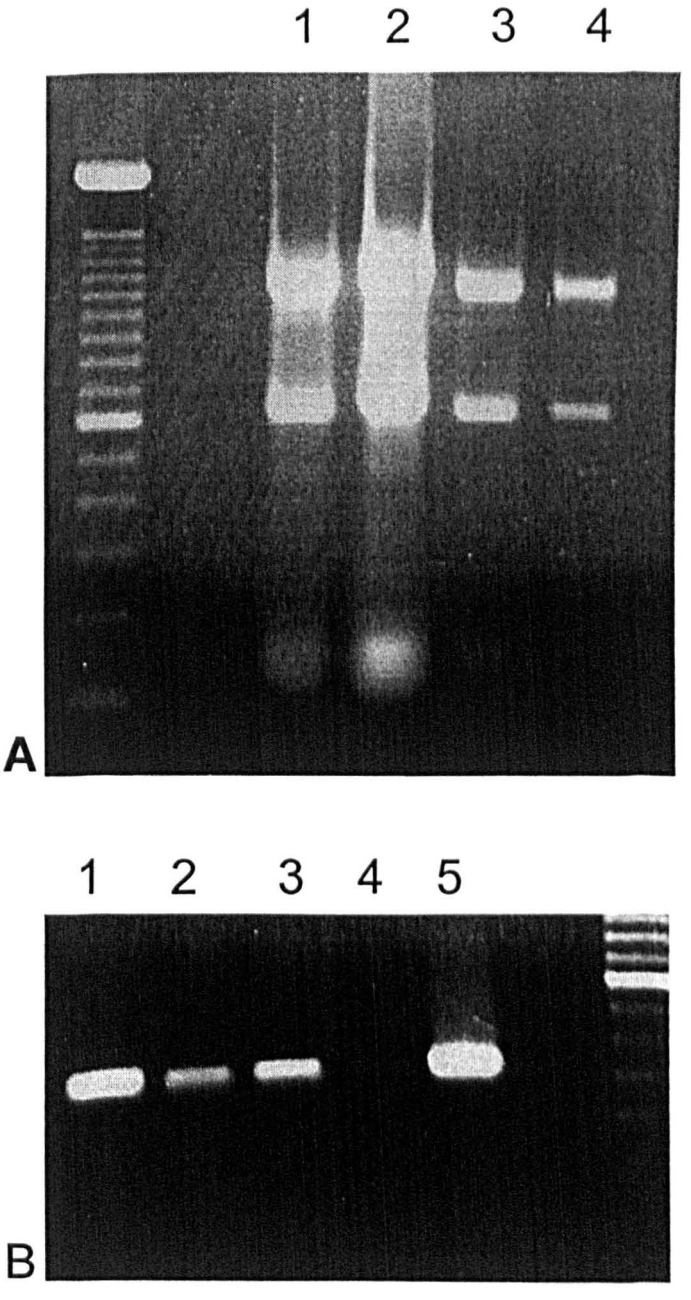


Figure 5.9: Agarose gel electrophoresis of A) RNA preparations from 11.5 dpc mouse embryo 1)heart, 2)whole, 3)forelimb 4)hindlimb; B) RT-PCR using mouse Tbx5 primers on RNA from 1)heart, 2)hindlimb, 3)negative control and 5)Plasmid DNA used as a positive control.

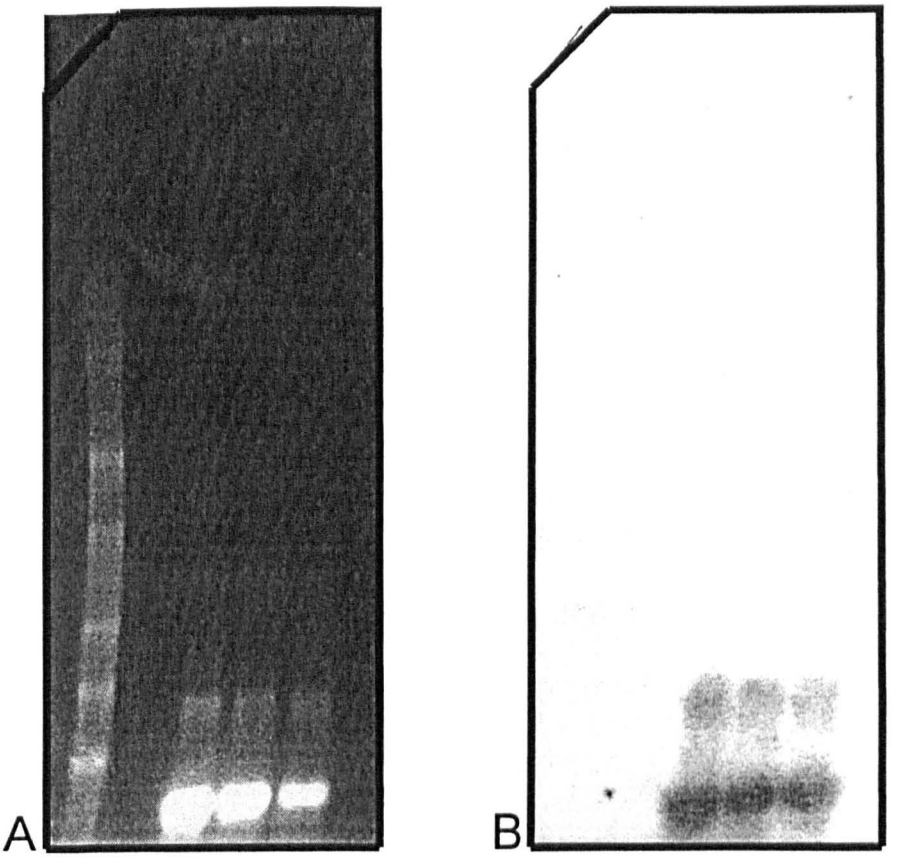


Figure 5.10: Northern blot hybridization experiment designed to test the conditions and reagents to be used in the expression analysis with RNA extracted from heart and limb mouse tissue in order to ensure its rational use. A) Agarose gel electrophoresis of different amounts of riboprobe synthesized with a mouse *Tbx5* cDNA clone. A RNA size marker was loaded in the leftmost lane. B) Autoradiography of the filter after the hybridization. The radioactive probe was produced by random priming of the same clone used to synthesize the riboprobe.

5.3.5.1 Clinical re-classification of patient 9/318

The work in region 4q was suspended when we learned that a member of the HOS15 family that was previously classified as not affected (9/318) showed minimal manifestations of the disease and was re-classified as affected.

As this patient showed two haplotypes in the 4q region not shared by any other affected subject, the assumption that the HOS2 locus is located within that interval is not compatible with the modification of the clinical data (see Figure 5.3).

As subject 9/318 was one of the cases labelled as "non-penetrant" assuming that the interval on chromosome 16 contains HOS2, the change in the clinical status classification reduced the number of non-penetrant cases to two, and according with an alternative interpretation (see below) to one (Figure 5.11). For that reason, further efforts were directed towards the region on chromosome 16.

5.3.5.2 Linkage intervals in the 16q region

The maximum interval was established using the method employed for the 4q region. Its centromeric limit was defined by the marker D16S3136, located at 49,260Kb from 16pter, centromeric to the selectin ligand

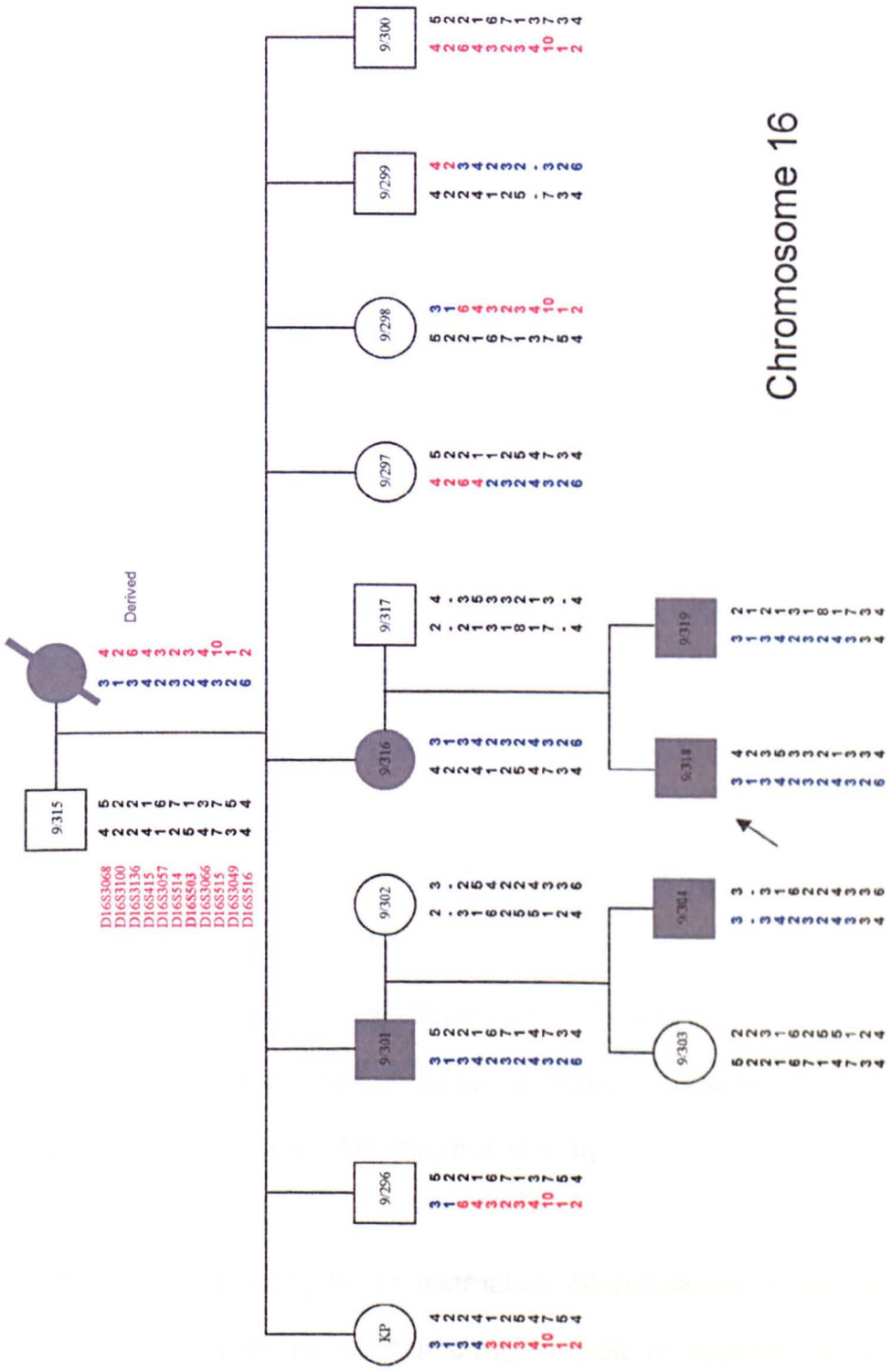


Figure 5.11: Haplotypes of markers in the HOS2 candidate region 16q. The haplotypes of the affected founder are derived. The "affected haplotype" is shown in blue. The arrow points to subject 9/318. Modified from Cross 2003.

interactor cytoplasmic-1 gene (*SLIC1*). The telomeric end of the maximum interval was defined by D16S3049, at 77,480kb, between LOC645957 and LOC729251.

The minimum interval was defined at the centromeric end by D16S3057, located at 56,087kb between the gene of the docking protein 4 (*DOK4*) and the coiled-coil domain containing 102A protein gene (*CCDC102A*). The telomeric limit was defined by D16S515, located in the intron 12 of the contactin associated protein-like 4 gene (*CNTNAP4*), at 75,070kb.

The length of the maximum interval was calculated as 28.220Mb, spanning most of the long arm of chromosome 16 and containing 297 gene related elements. The length of the minimum interval was calculated as 18.989Mb and comprises 214 elements.

Given the number of candidate genes in the minimum interval, it was considered impractical to adopt a screening strategy similar to that planned for the interval in chromosome 4q.

However, according to an alternative interpretation of the microsatellite typing data, a smaller candidate region could be defined between markers D16S3136 and D16S3057. Non-affected subject KP showed, in the chromosome inherited from her affected father an "affected haplotype" (shared by all affected subjects), corresponding to D16S3136 and the two markers centromeric to it, whereas the equally unaffected subject 9/297

showed in the paternal chromosome the affected haplotype in D16S3057 and seven markers telomeric to it (Figure 5.11). This suggests that a segment between D16S3136 and D16S3057 was the only part of the original affected paternal chromosome that was not transmitted to any of the two unaffected children and therefore liable to contain a hypothetical HOS2 locus.

This interpretation could also account for the unaffected status of subjects 9/296, 9/303, 9/297, 9/298, and 9/300 and the reclassification of 9/318 as affected, but not for subject 9/299, who being unaffected shows D16S3136 and D16S3057 alleles shared by all affected subjects. This patient could be considered as an example of non-penetrance if this interpretation is assumed as correct.

The chromosomal segment between D16S3136 and D16S3057 is 6.827Mb long and contains 74 gene-related elements. Amongst them, four were considered as good HOS2 candidates: *IRX3*, *IRX5*, *IRX6* and *SALL1* (Table 5.5).

5.3.5.3 The *Iroquois* B cluster

The *Iroquois* (*Irx*) genes were first discovered in Drosophila (*araucan*, *caupolican* and *mirror*), within a cluster in chromosome 3L (Gomez-Skarmeta et al. 1996). In mammals, an ancestral duplication event produced two very similar IRX clusters. In humans, the IrxA cluster is

Table 5.5 continues in next page..

D16S3136	Microsatellite
SLIC1	selectin ligand interactor cytoplasmic-1
CARD15	caspase recruitment domain family, member 15
CYLD	cylindromatosis (turban tumor syndrome)
LOC727992	hypothetical protein LOC727992
LOC643533	hypothetical protein LOC643533
LOC728654	hypothetical protein LOC728654
LOC643560	similar to Superoxide dismutase
SALL1	sal-like 1 (Drosophila)
UNGP1	uracil-DNA glycosylase pseudogene 1
LOC642659	heterogeneous nuclear ribonucleoprotein A1 pseudogene
LOC388276	hypothetical LOC388276
TNRC9	trinucleotide repeat containing 9
LOC643714	hypothetical protein LOC643714
LOC146253	tropomyosin-like
LOC390730	similar to prohibitin
CHD9	chromodomain helicase DNA binding protein 9
LOC643802	similar to M-phase phosphoprotein 10
RBL2	retinoblastoma-like 2 (p130)
FTS	fused toes homolog (mouse)
KIAA1005	KIAA1005 protein
FTO	fatso
IRX3	iroquois homeobox protein 3
LOC728792	hypothetical protein LOC728792
LOC643911	hypothetical protein LOC643911
LOC388279	hypothetical gene supported by AF275804
IRX5	iroquois homeobox protein 5
LOC654106	similar to iroquois homeobox protein 6
IRX6	iroquois homeobox protein 6
MMP2	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)
AYTL1	acyltransferase like 1
CAPNS2	calpain, small subunit 2
SLC6A2	solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2
LOC390732	similar to carboxylesterase 1 isoform c precursor
CES4	carboxylesterase 4-like
CES1	carboxylesterase 1 (monocyte/macrophage serine esterase 1)
CES7	carboxylesterase 7
GNAO1	guanine nucleotide binding protein (G protein), alpha activating activity polypeptide O
AMFR	autocrine motility factor receptor
NUDT21	nudix (nucleoside diphosphate linked moiety X)-type motif 21
OGFOD1	2-oxoglutarate and iron-dependent oxygenase domain containing 1
BBS2	Bardet-Biedl syndrome 2
MT4	metallothionein IV

MT3	metallothionein 3 (growth inhibitory factor (neurotrophic))
MT2A	metallothionein 2A
MT1L	metallothionein 1L (pseudogene)
MT1E	metallothionein 1E (functional)
MT1M	metallothionein 1M
MT1JP	metallothionein 1J (pseudogene)
MT1A	metallothionein 1A (functional)
MTM	metallothionein M
MT1B	metallothionein 1B (functional)
MT1F	metallothionein 1F (functional)
MT1G	metallothionein 1G
LOC727730	similar to metallothionein 1H-like protein
MT1H	metallothionein 1H
MT1X	metallothionein 1X
NUP93	nucleoporin 93kDa
MIRN138-2	microRNA 138-2
SLC12A3	solute carrier family 12 (sodium/chloride transporters), member 3
HERPUD1	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1
CETP	cholesteryl ester transfer protein, plasma
NOD27	nucleotide-binding oligomerization domains 27
CPNE2	copine II
NIP30	NEFA-interacting nuclear protein NIP30
RSPRY1	ring finger and SPRY domain containing 1
ARL2BP	ADP-ribosylation factor-like 2 binding protein
PLLP	plasma membrane proteolipid (plasmolipin)
CCL22	chemokine (C-C motif) ligand 22
CX3CL1	chemokine (C-X3-C motif) ligand 1
CCL17	chemokine (C-C motif) ligand 17
CIAPIN1	cytokine induced apoptosis inhibitor 1
COQ9	coenzyme Q9 homolog (S. cerevisiae)
POLR2C	polymerase (RNA) II (DNA directed) polypeptide C, 33kDa
DOK4	docking protein 4
D16S3057	Microsatellite

Table 5.5: Gene-related features in the region of chromosome 16q, between microsatelites D16S3136 and D16S3057. Genes screened for mutations in Family HOS15 are highlighted in grey.

located in chromosome 5p15 and includes the genes *IRX1*, *IRX2* and *IRX4*, whereas cluster B in chromosome 16q12 comprises *IRX3*, *IRX6* and *IRX6* (Ogura et al. 2001), all three included in the mutational analysis. Additionally, the developmentally expressed mohawk homeobox gene (*MKX*) highly related to *IRX3*, has been mapped to 10p12 (Anderson et al. 2006).

Interest in the genes of the *Iroquois* cluster B as potential HOS2 candidates was based in several observations. First, all three genes are highly expressed in the developing heart (Christoffels et al. 2000b; Mummenhoff et al. 2001). Second, a mouse deletion of the entire *Iroquois* B and the genes *Fto* and *Fta*, the mutation Fused toes (*Ft*), generated by transgenic insertional mutagenesis (van der Hoeven et al. 1994) produces, in the heterozygous animal, alterations of forelimb but not of the hindlimb, whereas the null embryos show other defects and die owing to heart malformations (Peters et al. 2002). Third, in a comparative gene expression analysis between mouse forelimb and hindlimb mouse tissues using serial analysis of gene expression (SAGE), expression of *Irx3* was observed in forelimb but not in hindlimb (Margulies et al. 2001). Fourth, a case of a human deletion 16q11.2-16q21, spanning, amongst other genes, *SALL1* (the gene mutated in Townes-Brocks syndrome, [TBS]) and the whole *Iroquois* B cluster, showed abnormalities typically associated with TBS plus radial aplasia (never found in patients with *SALL1* mutation), complex congenital heart defect and other malformations (Knoblauch et al. 2000).

5.3.5.4 SALL1

As mentioned above, mutations in *SALL1* cause Townes-Brocks syndrome (Kohlhase et al. 1998). This is an autosomal dominant disorder, typically characterized by malformations of the anus, hands, ears, sensorineural deafness, renal and genitourinary anomalies (Reid and Turner 1976; Townes and Brocks 1972). Foot defects, scoliosis, extra ribs, palsy of sixth and seventh cranial nerves and mental retardation have been also reported (Surka et al. 2001) and congenital heart defects are relatively rare (Monteiro de Pina-Neto 1984; Walpole and Hockey 1982). There is some phenotypic overlap with VATER (vertebral defects, anal atresia, tracheoesophageal fistula with esophageal atresia, and radial dysplasia) and VACTERL (with the addition of cardiac and limb abnormalities), two mostly sporadic associations that can be distinguished from TBS by the tracheal and oesophageal findings (Khoury et al. 1983).

SALL1 was included in the mutational analysis as phenotypic overlap occurs between Townes-Brocks and Holt-Oram syndromes. A sporadic case of a patient thought to have Holt-Oram syndrome showed absence of right thumb, aplasia of right radius, angulation of the right ulna, a lower dorsal hemivertebra, ventricular septal defect, duodenal atresia and rectovaginal fistula (Silver et al. 1972). The visceral abnormalities are not compatible with Holt-Oram syndrome and the right upper limb is more affected than the left; radial aplasia is not a feature of Townes-Brocks, and

the absence of tracheal or esophageal abnormalities is not compatible with VATER or VACTERL associations.

5.3.6 Mutational analysis

Four genes were screened in DNA samples of subjects from the HOS15 family: *IRX3* (four exons), *IRX5* (three exons), *IRX6* (six exons) and *SALL1* (three exons). In order to cover the all the coding regions of the genes, 26 amplicons were designed as described in Chapter 3 for the *MYH6* gene.

In most cases each exon and adjacent splicing regulatory regions were spanned by one amplicon. In the case of large exons (exons 1 and 2 in *IRX3*, exons 2 in *IRX5*, exons 1 and 5 in *IRX6*, and exon 3 in *SALL1*) two or more overlapping amplicons were designed to cover it completely. An exceptionally large exon (*SALL1*, exon 2) of 3.4kb was amplified in a long-PCR assay (amplicon 3.8kb) and subsequently sequenced using primers located approximately each 500 base pairs, as its size is not suitable for dHPLC analysis (Figures 5.12-5.15).

A pair of PCR primers was designed per amplicon (see Tables 5.6-5.9). The amplicons were screened for mutations either by direct sequence or dHPLC as described in Chapter 3. Because of their unusual GC content percentages (up to 80% for some regions exonic regions of *IRX3*, *IRX5*, and *SALL1*), several amplicons were impossible to amplify using conventional PCR methods. In these cases, amplification was achieved by

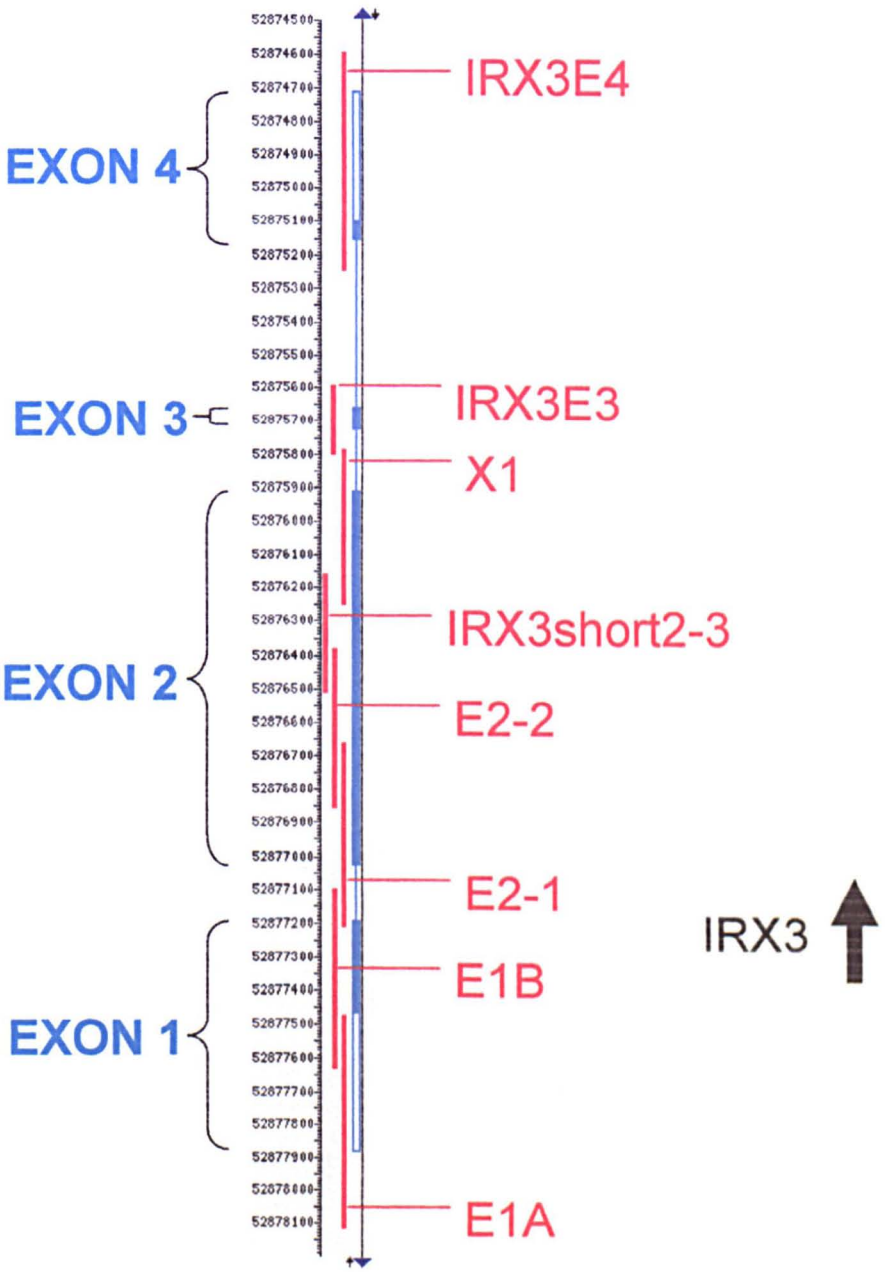


Figure 5.12: Scaled schematic representation of *IRX3* in 16q12.2. The orientation of the gene is telomeric to centromeric as indicated by the black arrow and here is drawn in its actual direction in the chromosome. The solid blue boxes represent the coding sequence of the exons, while empty blue boxes represent untranslated regions of the exons. The red lines indicate PCR amplicons used for the mutational analysis of the gene. The ruler indicate the distance in base-pairs from the telomere of the short arm of the chromosome.

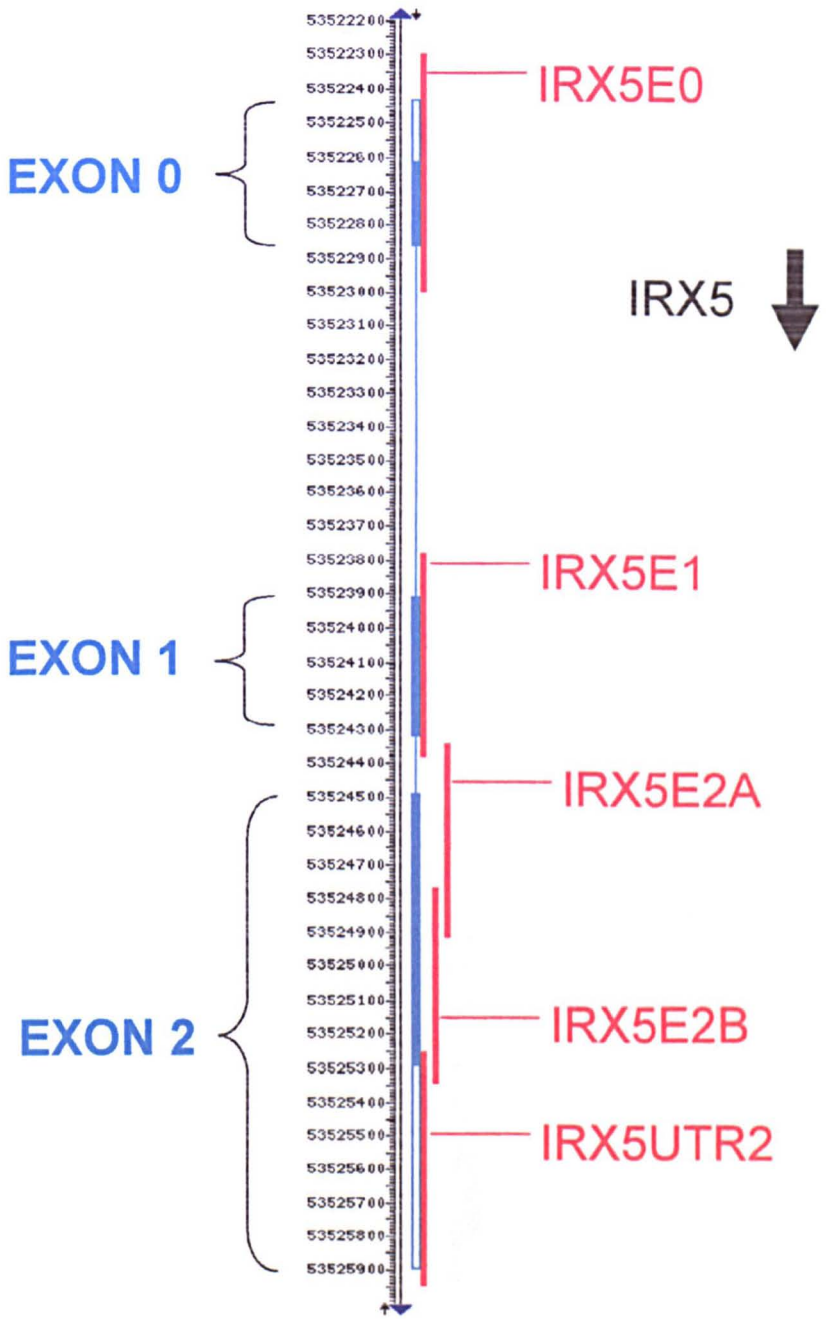


Figure 5.13: Scaled schematic representation of *IRX5* in 16q12.2. The orientation of the gene is centromeric to telomeric as indicated by the black arrow and here is drawn in its actual direction in the chromosome. The solid blue boxes represent the coding sequence of the exons, while empty blue boxes represent untranslated regions of the exons. The red lines indicate PCR amplicons used for the mutational analysis of the gene. The ruler indicate the distance in base-pairs from the telomere of the short arm of the chromosome.

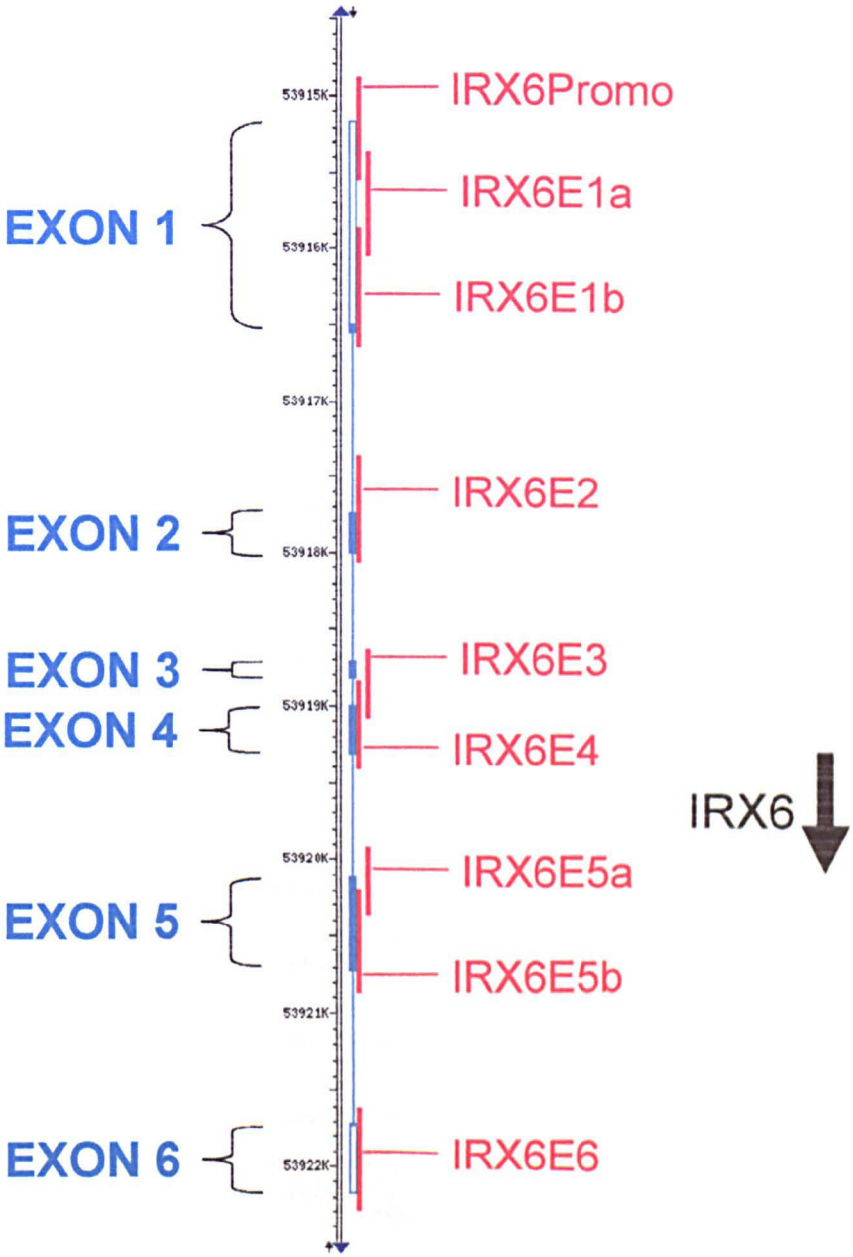


Figure 5.14: Scaled schematic representation of *IRX6* in 16q12.2. The orientation of the gene is centromeric to telomeric as indicated by the black arrow and here is drawn in its actual direction in the chromosome. The solid blue boxes represent the coding sequence of the exons, while empty blue boxes represent untranslated regions of the exons. The red lines indicate PCR amplicons used for the mutational analysis of the gene. The ruler indicate the distance in kilobases from the telomere of the short arm of the chromosome.

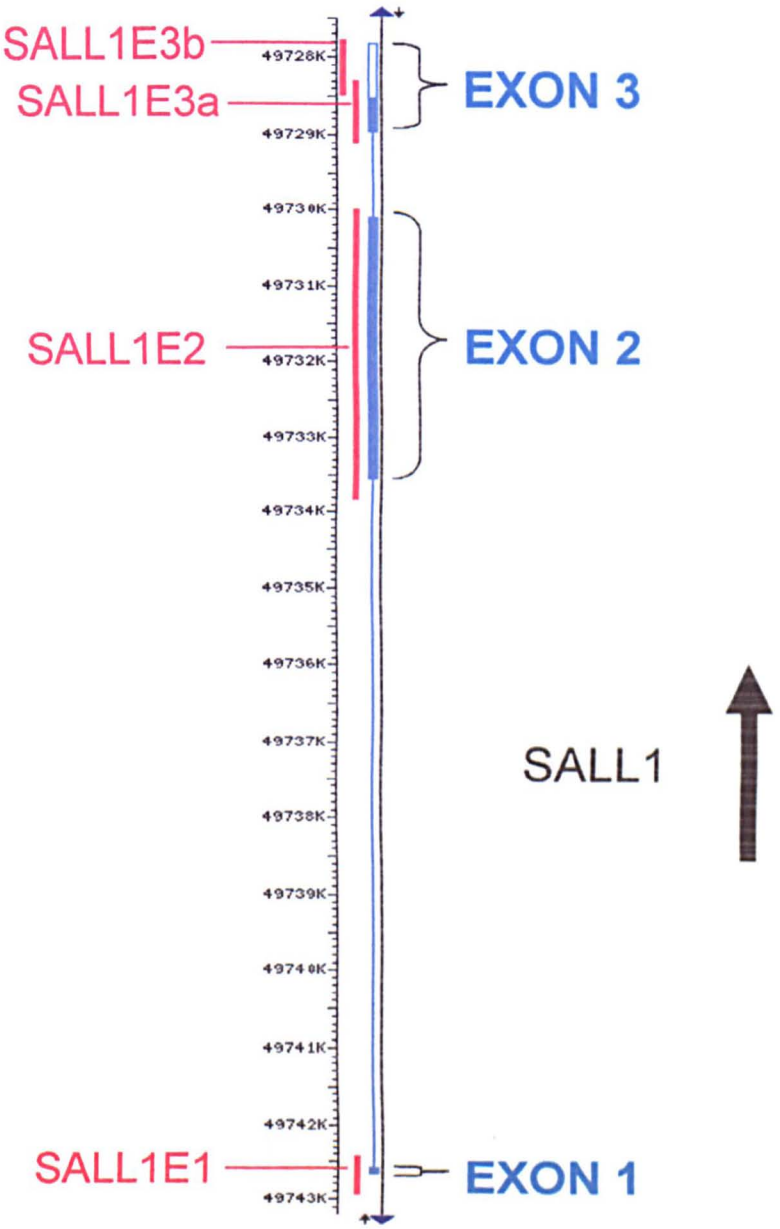


Figure 5.15: Scaled schematic representation of SALL1 in 16q12.1. The orientation of the gene is telomeric to centromeric as indicated by the black arrow and here is drawn in its actual direction in the chromosome. The solid blue boxes represent the coding sequence of the exons, while empty blue boxes represent untranslated regions of the exons. The red lines indicate PCR amplicons used for the mutational analysis of the gene. The ruler indicate the distance in kilobases from the telomere of the short arm of the chromosome. The large exon 2 was analyzed as a single long-PCR, and sequenced using six internal primers and a smaller, nested PCR amplification.

addition of DMSO as well as the proprietary reagent GC-melt (BD Biosciences) to the PCR reaction. PCR products obtained by this method could not be processed by dHPLC as proper absorption and elution of the mixture by the chromatography column is prevented by these compounds.

Table 5.6: Primers used for mutational analysis of *IRX3*	
Primer	Sequence
E1AF	tggaaaggtcgcggggagtatcg
E1AR	cggggcacggacggagagg
E1BF	ccgccgcggagcagatcaatagg
E1BR	tcctggcctgcacccctctagtcc
Exon2-1F	atcttcccgcagctggtaagagccc
Exon2-1R	cgctcccataagcgtttccctcctc
Exon2-2F	agcaccgcaagaacccctaccccacc
Exon2-2R	cctctaagccctcagagctatcttc
IRX3short2-3F	cggccaccgagcctgagctgtcc
IRX3short2-3R	gaccgctgcccccggtggagacc
X1F	atctggtccctcgcggagactg
X1R	gcagaaagcaggagtggagagg
IRX3E3F	tctccactcctgctttctgc
IRX3E3R	ctcggcgtcctctccttt
IRX3E4F	cgatcggccccaatccaagtagg
IRX3E4R	actcggtcccgattcgtctctcg

Table 5.7: Primers used for mutational analysis of *IRX5*

Primer	Sequence
IRX5E0F	gcaaaggcaaaagcagagc
IRX5E0R	aatcgcccaagtttgaagg
IRX51F	cccgtaggaagctggagtg
IRX5E2F	ggcgggagtaaaaaggaaaa
IRX51R	ctccaggcctctcttttcct
IRX5E2BF	ggccctcggttatccatt
IRX5E2AR	atggagagccctcdttcc
IRX5UTR2F	cctatgaattgaagaaaggtatgtcc
IRX52R	ctgccaaggccatgttttta
IRX5UTR2R	tctgtggaacctttcaatcc

Table 5.8: Primers used for mutational analysis of *IRX6*	
Primer	Sequence
IRX6PromoF	aaattggaggttccatgtctcg
IRX6E1aF	agagaagctccaaggtcaagg
IRX6PromoR	gtaaagtttgttcacgccacagg
IRX6E1bF	gaggcgtttctcctacttctcc
IRX6E1aR	ctcacccacactcacacttgc
IRX6E1bR	acaagaggaggatgatcagagg
IRX6E2F	ggtctctaaggccaccttctagc
IRX6E2R	ctggatttcacctgtctctactgc
IRX6E3F	ttgaaccagacccttagctacca
IRX6E4F	gactggcactgtgagtctttcc
IRX63R	ccttgagtgtactggtggtctcc
IRX6E4R	ggaccttccagaaaatctaccc
IRX6E5aF	agaaattctttcccaaggacagg
IRX6E5bF	ctgaagacgaggaggtagtgg
IRX6E5aR	gaagggtcattgaaggagaagc
IRX6E5bR	gagcgctgtaaagcaatctcc
IRX6E6F	cttgccttcagaagagcatgg
IRX6E6R	ccagctactaaaggaactgagg

Table 5.9: Primers used for mutational analysis of *SALL1*

Primer	Sequence
SALL1E1F	caaatcacgaactaattgattaagg
SALL1E1R	attacgccgagtggaagaagc
SALL1E2F	catcatgcaccaatgtctcc
SALL1EgapF	gttcacgaacgctgtggtc
SALL1EgapR	gccacaaatgtcacaagcag
SALL1E2R	atctgggctgatgactctgg
SALL1E3aF	gaatagacatttaggccacttgc
SALL1E3bF	ctatggtggcctcctactcc
SALL1E3aR	tttgcaaagcaaggttatatcg
SALL1E3bR	tgttacgtctcagtttcaactacc

In the case of amplicons that did not required extra reagents, the sequence of each individual amplicon was analyzed using the Navigator program, part of the dHPLC Wave system software (Transgenomic) in order to calculate the optimum temperatures for mutation detection.

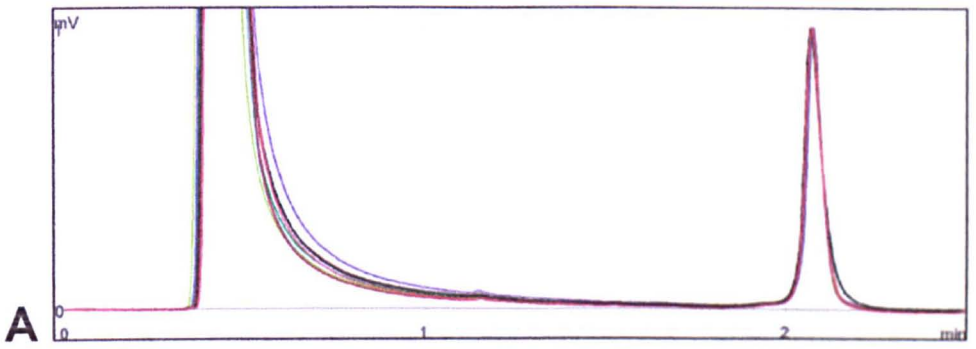
dHPLC was carried out using the WAVE System (Transgenomic), as described in Chapter 3.

5.3.7 Results of mutational analysis

Good quality dHPLC traces were obtained for all the amplicons that did not require the addition of DMSO or GC-Melt for PCR amplification. No sequence variants segregating with the disease were detected. Some representative traces are shown in Figures 5.16 and 5.17.

Good quality sequence data were obtained for all GC-rich amplicons and the 3.8Kb *SALL1* exon 2 PCR amplicon.

A three base pair deletion was detected with exon 2 of the *SALL1* gene (Figure 5.18). This was an in-frame deletion of a serine codon in the 164 position. This variant was not considered relevant for the phenotype as it has been reported as a polymorphism (Kohlhase 2000) and does not segregate with the disease in the HOS15 family.



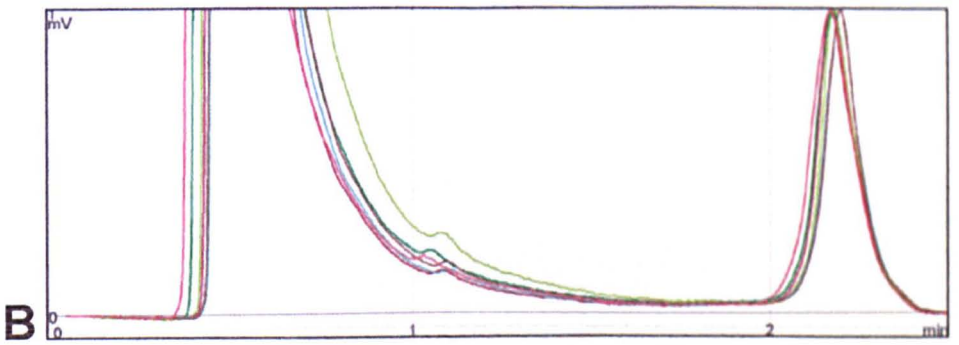
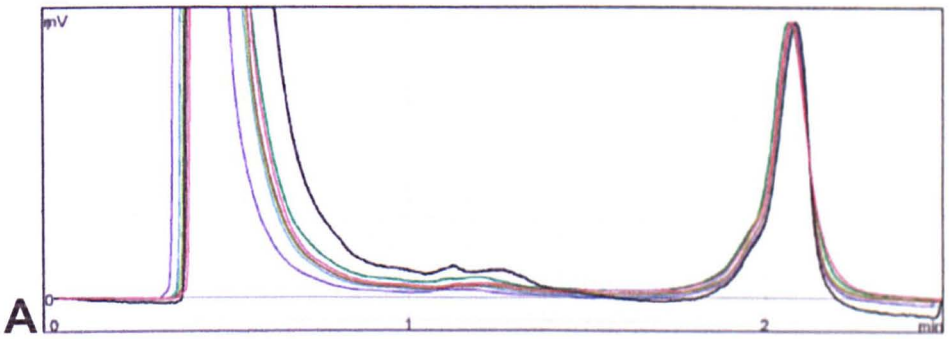


Figure 5.16: Two representative examples of dHPLC traces obtained during the mutational analysis of the selected candidate genes in the 16q region. A) IRX3 exon 4; B) IRX5 exon 1.



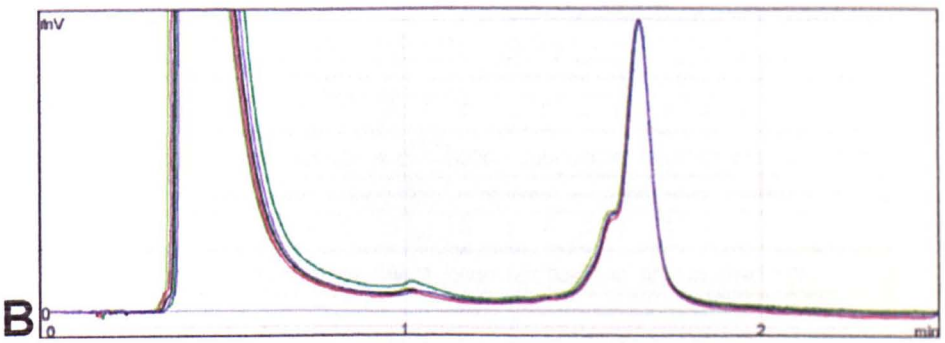


Figure 5.17: Two representative examples of dHPLC traces obtained during the mutational analysis of the selected candidate genes in the 16q region. A) IRX 6 exon 4; B) SALL1 exon 3, amplicon a.

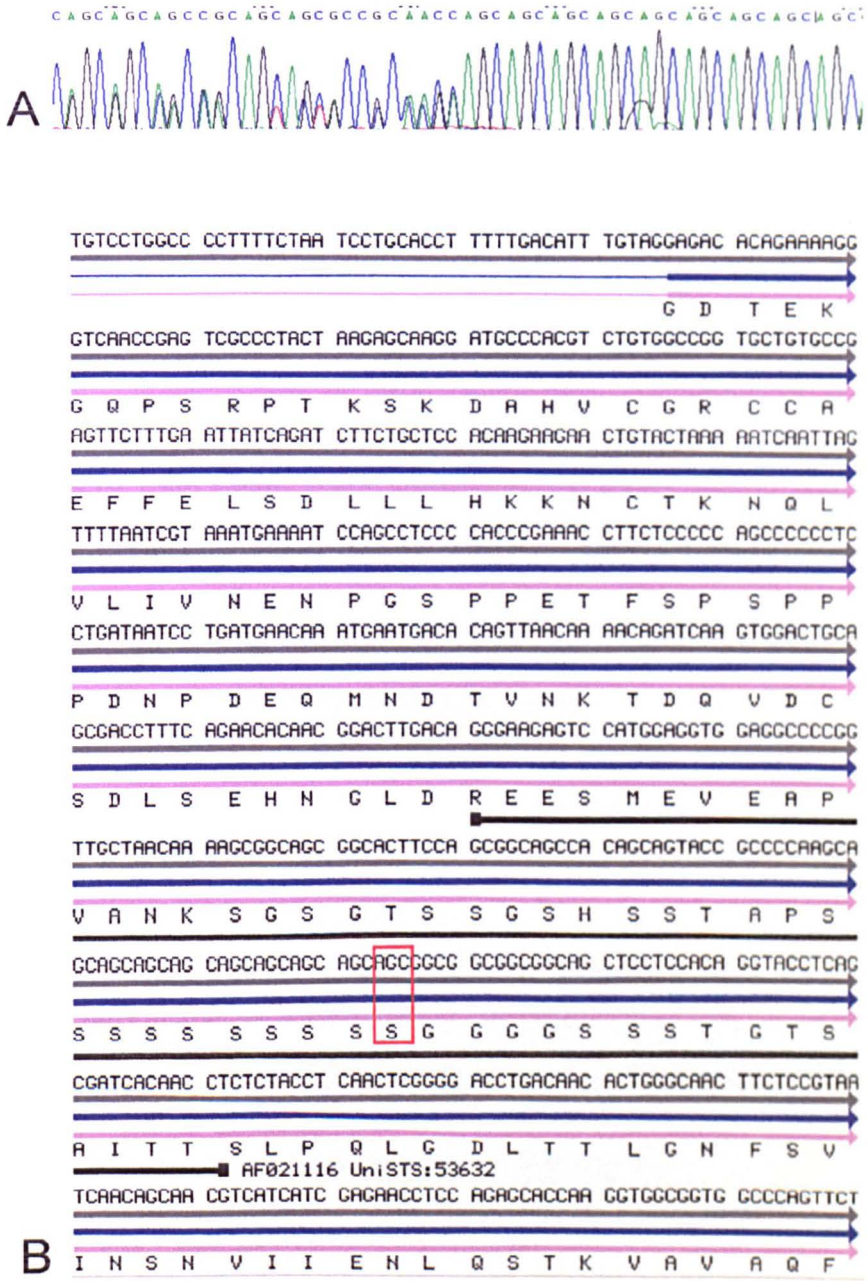


Figure 5.18: In-frame three base pair deletion found in subjects from the HOS15 family corresponding the exon 2 of *SALL1*. The product of this variant allele lacks of one of a long tract of serine residues. This variant has been described as a benign polymorphism and does not segregate with the disease. A) Electropherogram showing the position of the deletion (reverse and complement trace); B)The deletion in its annotated sequence context.

No other coding variant was identified in the remaining direct sequence screened PCR amplicons.

A control probe included in the *MYH6* MAPH probe set (Chapter 4) is specific for a short segment of the *IRX3* gene. The second objective of this probe was to detect complete deletions of *IRX3* when DNA samples from family HOS15 individuals were processed with the *MYH6* MAPH probe set. This of course was not intended as a comprehensive *IRX3* deletion detection strategy but if positive it could provide important information. The HOS15 samples showed a normal MAPH trace.

5.4 Discussion

Holt-Oram syndrome is a genetically heterogeneous entity. Mutations in *TBX5* and *SALL4* have been found in sporadic and familial cases. Some affected families do not show linkage to any of them. In one such family (HOS15), a previous genome-wide linked analysis identified a candidate region in chromosome 4q.

It has been estimated that a LOD score higher than 3 can be used as evidence of linkage it two point mapping with a 5% chance of error (Morton 1955), whereas a threshold of 3.3 can be used for genome-wide linkage analysis (Lander and Schork 1994).

Although according to simulations with the program SLINK (Ott 1976), was estimated that, given the size of family HOS15, the maximum LOD score obtainable by linkage would be of 2.25, and that the highest actual LOD score obtained during the fine mapping stage was 1.75 for the D4S3046 marker (Cross 2003), it was considered that the available data should at least allow meaningful haplotype analysis to define candidate regions (probably large) with an acceptable degree of confidence. It was as well considered that the complexities of a large number of candidate genes could be overcome with an efficient gene expression screening as we are dealing with a purely developmental trait.

This chapter summarizes further efforts done to screen the genes located within the 4q linkage interval and subsequently another interval in 16q, once modifications in the clinical classification of a family member rendered the former incompatible with the new data.

The second region considered in 16q, however, contains a much larger number of candidates and the expression screening was considered impractical.

Mutational analysis was carried out in *IRX3*, *IRX5*, *IRX6*, and *SALL1*, all four located in the 16q interval. No sequence *variant compatible* with a disease causing mutation was found.

5.4.1 Conserved non-coding sequences in the *Iroquois* B cluster

The analysis of the sequences provided by the Human Genome Project has highlighted, besides coding sequences, certain elements that are not likely to produce a functional transcript and that have been conserved during evolution. This "conserved non-genic sequences" or "conserved non-coding sequences" (CNS) are thought to have an important role in regulation of gene expression and some times they cluster in regions with very little gene content or "gene deserts" (Dermitzakis et al. 2005).

The 5Mb gene-poor region that contains the *SALL1*, and the *Iroquois* cluster B is located within the largest block of unbroken synteny across human/mouse/dog/chicken found in chromosome 16, where approximately 59% of all chromosome 16 human/mouse/fugu CNSs are clustered (Martin et al. 2004)

In transgenesis experiments using Zebrafish and *Xenopus tropicalis* it has been noticed that the intergenic regions between the *Iroquois* cluster B genes function as modular tissue or organ specific enhancers, where specific segments control the expression of a reporter gene in specific tissues and developmental stages (de la Calle-Mustienes et al. 2005).

Even though no heterozygous potentially pathogenic sequence variation was found in the coding regions of the *Iroquois* B, mutations in the non-coding regions of the cluster could be relevant to the phenotype.

5.4.2 Limitations of the HOS15 family for linkage

The main problem facing the identification of the mutation responsible for the phenotype in family HOS15 is the great length of the linkage intervals which is a direct consequence of the small size of the family. If more informative meioses are studied, should new members of the youngest generation or members of a new generation be born, higher LOD scores could be obtained.

5.4.3 Phenotypic variability and future mapping efforts

A striking phenotypic variability has been observed between Tbx5 knock-out mice. Heterozygous Black-Swiss mice showed mild skeletal abnormalities and mild cardiac malformations whereas 129SvEv mice displayed complex cardiac malformations and died *in utero*. This points to an important role of modifier genes in the establishment of heart abnormalities in the phenotype associated with *Tbx5* mutation (Bruneau et al. 2001). If the effect of modifier genes is even greater when other hypothetical Holt-Oram causing genes are mutated, incomplete penetrance could pose further difficulties to their identification.

Chapter 6 FINAL DISCUSSION

Congenital heart defects (CHD) belong to a heterogeneous group of diseases of complex aetiology. Most of them appear sporadically in the population, probably because of the interaction between an unknown number of genetic determinants and environmental factors.

Even though the discovery of a new gene responsible for a Mendelian form of a particular type of CHD has little immediate impact over the way this kind of diseases can be prevented or treated, the knowledge gained in that way is a useful tool in the dissection of the mechanisms of cardiac development, an extremely complex process.

The discovery of a mutation of the gene *MYH6* as a cause of a Mendelian form of atrial septal defect (Ching 2005), the first gene encoding a structural protein to be related to this particular phenotype, inspired the first main topic of this thesis: The search for point mutations and copy number variations in the *MYH6* gene in subjects with apparently sporadic CHDs.

No *de novo* mutations were discovered in our CHD cohort. Several inherited mutations were identified. In particular a nonsense mutation and

a *splicing* acceptor site mutation inherited from healthy parents are a striking feature of these findings that deserve *further investigation.* Possible explanations were developed in the discussion of Chapter 3.

No copy number variations were discovered. Recurrent deletions are a cause of several Mendelian diseases mostly due to non-allelic homologous recombination events between very similar genes or repetitive elements.

MYH6 and *MYH7* are remarkably similar and are located in tandem in the long arm of chromosome 14. A previously reported non-allelic homologous recombination event made the prospect of a similar event being responsible for a small number of CHD cases seemed very interesting.

The second main topic in this work is related to a purely Mendelian condition: Holt-Oram syndrome.

The main Holt-Oram gene (*TBX5* in chromosome 12) was identified a decade ago and numerous advances in developmental biology a molecular pathology have been achieved as a direct consequence. The identification of a new gene (*SALL4* in chromosome 20) capable of producing a phenotype clinically indistinguishable from Holt-Oram syndrome and of families without linkage with either *TBX5* or *SALL4* was very encouraging. Nevertheless, the size of the biggest family in this case is not large enough to provide a chromosomal interval small enough to

conduct a comprehensive and efficient mutational screening. The growing of this family and the identification of new families not linked to chromosomes 12 or 20 should bring new hope.

BIBLIOGRAPHY

Akrami SM, Winter RM, Brook JD, Armour JA (2001) Detection of a large TBX5 deletion in a family with Holt-Oram syndrome. J Med Genet 38: E44

Alexander J, Rothenberg M, Henry GL, Stainier DY (1999) casanova plays an early and essential role in endoderm formation in zebrafish. Dev Biol 215: 343-57

Alsan BH, Schultheiss TM (2002) Regulation of avian cardiogenesis by Fgf8 signaling. Development 129: 1935-43

Alvarez AD, Shi W, Wilson BA, Skeath JB (2003) pannier and pointedP2 act sequentially to regulate Drosophila heart development. Development 130: 3015-26

Anan R, Greve G, Thierfelder L, Watkins H, McKenna WJ, Solomon S, Vecchio C, Shono H, Nakao S, Tanaka H, et al. (1994) Prognostic implications of novel beta cardiac myosin heavy chain gene mutations that cause familial hypertrophic cardiomyopathy. J Clin Invest 93: 280-5

Anderson DM, Arredondo J, Hahn K, Valente G, Martin JF, Wilson-Rawls J, Rawls A (2006) Mohawk is a novel homeobox gene expressed in the developing mouse embryo. Dev Dyn 235: 792-801

Anderson RH, Webb S, Brown NA, Lamers W, Moorman A (2003a) Development of the heart: (2) Septation of the atriums and ventricles. Heart 89: 949-58

Anderson RH, Webb S, Brown NA, Lamers W, Moorman A (2003b) Development of the heart: (3) formation of the ventricular outflow tracts, arterial valves, and intrapericardial arterial trunks. Heart 89: 1110-8

Armour JA, Sismani C, Patsalis PC, Cross G (2000) *Measurement of locus copy* number by hybridisation with amplifiable probes. Nucleic Acids Res 28: 605-9

Azpiazu N, Frasch M (1993) tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of Drosophila. Genes Dev 7: 1325-40

Bamford RN, Roessler E, Burdine RD, Saplakoglu U, dela Cruz J, Splitt M, Towbin J, Bowers P, Marino B, Schier AF, Shen MM, Muenke M, Casey B (2000) Loss-of-function mutations in the EGF-CFC gene CFC1 are associated with human left-right laterality defects. Nature Genetics 26: 365-369

Bao ZZ, Bruneau BG, Seidman JG, Seidman CE, Cepko CL (1999) Regulation of chamber-specific gene expression in the developing heart by Irx4. Science 283: 1161-1164

Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soults J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, Renault B, Kucherlapati R, Seidman JG, Seidman CE (1997) Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome [published erratum appears in Nat Genet 1997 Apr;15(4):411]. Nat Genet 15: 30-5

Basson CT, Cowley GS, Solomon SD, Weissman B, Pozanski AK, Traill TA, Seidman JG, Seidman CE (1994) The clinical and genetic spectrum of the Holt-Oram syndrome(heart-hand syndrome I). New Engl. J. Med. 330: 885-891

Basson CT, Huang T, Lin RC, Bachinsky DR, Weremowicz S, Vaglio A, Bruzzone R, Quadrelli R, Lerone M, Romeo G, Silengo M, Pereira A, Krieger J, Mesquita SF, Kamisago M, Morton CC, Pierpont ME, Muller CW, Seidman JG, Seidman CE (1999) Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. Proc Natl Acad Sci U S A 96: 2919-24

Berdougo E, Coleman H, Lee DH, Stainier DY, Yelon D (2003) Mutation of weak atrium/atrial myosin heavy chain disrupts atrial function and influences ventricular morphogenesis in zebrafish. Development 130: 6121-9

Berg JS, Powell BC, Cheney RE (2001) A millennial myosin census. Mol Biol Cell 12: 780-94

Biben C, Harvey, R. P. (1997) Homeodomain factor *Nkx2-5 controls left/right* asymmetric expression of bHLH gene eHand during murine heart development. Genes & Development 11: 1357-1369

Bleyl SB, Mumford BR, Thompson V, Carey JC, Pysher TJ, Chin TK, Ward K (1997) Neonatal, lethal noncompaction of the left ventricular myocardium is allelic with Barth syndrome. Am J Hum Genet 61: 868-72

Bodmer R, . Frasch, M. (1999) Genetic determination of drosophila heart development. In: Harvey RP, Rosenthal, N. (ed) In: Heart development. Academic press, San Diego

Bonnet D, Pelet A, Legeai-Mallet L, Sidi D, Mathieu M, Parent P, Plauchu H, Serville F, Schinzel A, Weissenbach J, Kachaner J, Munnich A, Lyonnet S (1994) A gene for Holt-Oram syndrome maps to the distal long arm of chromosome 12. Nature Genet. 6: 405-408

Boughman JA (1993) The genetics of congenital heart disease. In: Ferencz C (ed) Epidemyology of congenital heart disease: the Baltimore-Washington infant heart study 1981-1989. Futura Mout Kisco, NY, pp 123-164

Boughman JA, Berg KA, Astemborski JA, Clark EB, McCarter RJ, Rubin JD, Ferencz C (1987) Familial risks of congenital heart defect assessed in a population-based epidemiologic study. Am J Med Genet 26: 839-49

Bourguignon LY, Zhu H, Zhou B, Diedrich F, Singleton PA, Hung MC (2001) Hyaluronan promotes CD44v3-Vav2 interaction with Grb2-p185(HER2) and induces Rac1 and Ras signaling during ovarian tumor cell migration and growth. J Biol Chem 276: 48679-92

Brand T (2003) Heart development: molecular insights into cardiac specification and early morphogenesis. Dev Biol 258: 1-19

Brassington AM, Sung SS, Toydemir RM, Le T, Roeder AD, Rutherford AE, Whitby FG, Jorde LB, Bamshad MJ (2003) Expressivity of Holt-Oram syndrome is not predicted by TBX5 genotype. Am J Hum Genet 73: 74-85

Brown CB, Boyer AS, Runyan RB, Barnett JV (1999) Requirement of type III TGF-beta receptor for endocardial cell transformation in the heart. Science 283: 2080-2082

Brown CB, Wenning JM, Lu MM, Epstein DJ, Meyers EN, Epstein JA (2004) Cre-mediated excision of Fgf8 in the Tbx1 expression domain reveals a critical role for Fgf8 in cardiovascular development in the mouse. Dev Biol 267: 190-202

Bruneau BG, Bao ZZ, Tanaka M, Schott JJ, Izumo S, Cepko CL, Seidman JG, Seidman CE *(2000) Cardiac expression* of the ventricle-specific homeobox gene Irx4 is modulated by Nkx2-5 and *dHand. Developmental Biology* 217: 266-277

Bruneau BG, Logan M, Davis N, Levi T, Tabin CJ, Seidman JG, Seidman CE (1999) Chamber-specific cardiac expression of Tbx5 and heart defects in Holt- Oram syndrome. Dev Biol 211: 100-8.

Bruneau BG, Nemer G, Schmitt JP, Charron F, Robitaille L, Caron S, Conner DA, Gessler M, Nemer M, Seidman CE, Seidman JG (2001) A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. Cell 106: 709-21

Burn J (2002) The aetiology of congenital heart disease. In: Anderson RH (ed) Paediatric cardiology, vol 1. Churchill Livingston, London

Burn J, Brennan P, Little J, Holloway S, Coffey R, Somerville J, Dennis NR, Allan L, Arnold R, Deanfield JE, Godman M, Houston A, Keeton B, Oakley C, Scott O, Silove E, Wilkinson J, Pembrey M, Hunter AS (1998) Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. Lancet 351: 311-6

Burn J, Corney G (1984) Congenital heart defects and twinning. Acta Genet Med Gemellol (Roma) 33: 61-9

Burrows NP, Nicholls AC, Richards AJ, Luccarini C, Harrison JB, Yates JR, Pope FM (1998) A point mutation in an intronic branch site results in aberrant splicing of COL5A1 and in Ehlers-Danlos syndrome type II in two British families. Am J Hum Genet 63: 390-8

Cai CL, Liang X, Shi Y, Chu PH, Pfaff SL, Chen J, Evans S (2003) Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. Dev Cell 5: 877-89

Camenisch TD, Spicer AP, Brehm-Gibson T, Biesterfeldt J, Augustine ML, Calabro A, Jr., Kubalak S, Klewer SE, McDonald JA (2000) Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. J Clin Invest 106: 349-60

Campione M, Ros MA, Icardo JM, Piedra E, Christoffels VM, Schweickert A, Blum M, Franco D, Moorman AF (2001) Pitx2 expression defines a left cardiac lineage of cells: evidence for atrial and ventricular molecular isomerism in the iv/iv mice. Dev Biol 231: 252-64

Campos-Ortega JA, Hartenstein, V. (1985) The Embryonic development of Drosophila melanogaster. Springer-Verlag, Berlin

Capdevila J, Vogan KJ, Tabin CJ, Belmonte JCI (2000) Mechanisms of left-right determination in vertebrates. Cell 101: 9-21

Carlson BM (2004) Human Embryology and developmental biology. Mosby, *Philadelphi*

Carmeliet P, Ferreira V, Breier G, Pollefeyt S, *Kieckens L, Gertsenstein M,* Fahrig M, Vandenhoeck A, Harpal K, Eberhardt C, Declercq C, Pawling J, Moons L, Collen D, Risau W, Nagy A (1996) Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. Nature 380: 435-439

Carrasquillo MM, McCallion AS, Puffenberger EG, Kashuk CS, Nouri N, Chakravarti A (2002) Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat Genet 32: 237-44

Chen CY, Croissant J, Majesky M, Topouzis S, McQuinn T, Frankovsky MJ, Schwartz RJ (1996) Activation of the cardiac alpha-actin promoter depends upon serum response factor, Tinman homologue, Nkx-2.5, and intact serum response elements. Dev Genet 19: 119-30

Chen JN, vanEeden FJM, Warren KS, Chin A, NussleinVolhard C, Haffter P, Fishman MC (*1997) Left-right pattern* of cardiac BMP4 may drive asymmetry of the heart in zebrafish. Development *124: 4373-4382*

Cheng SK, Olale F, Bennett JT, Brivanlou AH, Schier AF (2003) EGF-CFC proteins are essential coreceptors for the TGF-beta signals Vg1 and GDF1. Genes Dev 17: 31-6

Ching YH (2001) Molecular genetics of human atrial septal defects, University of Nottingham

Ching YH, Ghosh TK, Cross SJ, Packham EA, Honeyman L, Loughna S, Robinson TE, Dearlove AM, Ribas G, Bonser AJ, Thomas NR, Scotter AJ, Caves LS, Tyrrell GP, Newbury-Ecob RA, Munnich A, Bonnet D, Brook JD (2005) Mutation in myosin heavy chain 6 causes atrial septal defect. Nat Genet 37: 423-8

Christoffels VM, Habets PE, Franco D, Campione M, de Jong F, Lamers WH, Bao ZZ, Palmer S, Biben C, Harvey RP, Moorman AF (2000a) Chamber formation and morphogenesis in the developing mammalian heart. Dev Biol 223: 266-78

Christoffels VM, Keijser AG, Houweling AC, Clout DE, Moorman AF (2000b) Patterning the embryonic heart: identification of five mouse Iroquois homeobox genes in the developing heart. Dev Biol 224: 263-74

Ciruna B, Rossant J (2001) FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. Dev Cell 1: 37-49

Conrad DF, Andrews TD, Carter NP, Hurles ME, Pritchard JK (2006) A high-resolution survey of deletion polymorphism in the human genome. Nat Genet 38: 75-81

Cross S (2003) Genetics of congenital heart disease and Holt-Oram syndrome., University of Nottingham

Cross SJ, Ching YH, Li QY, Armstrong-Buisseret L, Spranger S, Lyonnet S, Bonnet D, Penttinen M, Jonveaux P, Leheup B, Mortier G, Van Ravenswaaij C, Gardiner CA (2000) The mutation spectrum in Holt-Oram syndrome. J Med Genet 37: 785-7.

Curran ME, Atkinson DL, Ewart AK, Morris CA, Leppert MF, Keating MT (1993) The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis. Cell 73: 159-68

Czeizel A, Pornoi A, Peterffy E, Tarcal E (1982) Study of children of parents operated on for congenital cardiovascular malformations. Br Heart J 47: 290-3

D'Souza SW, Robertson IG, Donnai D, Mawer G (1991) Fetal phenytoin exposure, hypoplastic nails, and jitteriness. Arch Dis Child 66: 320-4

Dasgupta C, Martinez AM, Zuppan CW, Shah MM, Bailey LL, Fletcher WH (2001) Identification of connexin43 (alpha1) gap junction gene mutations in patients with hypoplastic left heart syndrome by denaturing gradient gel electrophoresis (DGGE). Mutat Res 479: 173-86

Davidson B, Levine M (2003) Evolutionary origins of the vertebrate heart: Specification of the cardiac lineage in Ciona intestinalis. Proc Natl Acad Sci U S A 100: 11469-73

Davis DL, Edwards AV, Juraszek AL, Phelps A, Wessels A, Burch JB (2001) A GATA-6 gene heart-region-specific enhancer provides a novel means to mark and probe a discrete component of the mouse cardiac conduction system. Mech Dev 108: 105-19

de la Calle-Mustienes E, Feijoo CG, Manzanares M, Tena JJ, Rodriguez-Seguel E, Letizia A, Allende ML, Gomez-Skarmeta JL (2005) A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. Genome Res 15: 1061-72

de la Pompa JL, Timmerman LA, Takimoto H, Yoshida H, Elia AJ, Samper E, Potter J, Wakeham A, Marengere L, Langille BL, Crabtree GR, Mak TW (1998) Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. Nature 392: 182-186

Deininger PL, Batzer MA (1999) Alu repeats and human disease. Mol Genet Metab 67: 183-93

Dentice M, Morisco C, Vitale M, Rossi G, Fenzi G, Salvatore D (2003) The different cardiac expression of the type 2 iodothyronine deiodinase gene between human and rat is related to the differential response of the Dio2 genes to Nkx-2.5 and GATA-4 transcription factors. Mol Endocrinol 17: 1508-21

Dermitzakis ET, Reymond A, Antonarakis SE (2005) Conserved non-genic sequences - an unexpected feature of mammalian genomes. Nat Rev Genet 6: 151-7

Dodou E, Verzi MP, Anderson JP, Xu SM, Black BL (2004) Mef2c is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. Development 131: 3931-42

Donaudy F, Ferrara A, Esposito L, Hertzano R, Ben-David O, Bell RE, Melchionda S, Zelante L, Avraham KB, Gasparini P (2003) Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. Am J Hum Genet 72: 1571-7

Donaudy F, Snoeckx R, Pfister M, Zenner HP, Blin N, Di Stazio M, Ferrara A, Lanzara C, Ficarella R, Declau F, Pusch CM, Nurnberg P, Melchionda S, Zelante L, Ballana E, Estivill X, Van Camp G, Gasparini P, Savoia A (2004) Nonmuscle myosin heavy-chain gene MYH14 is expressed in cochlea and mutated in patients affected by autosomal dominant hearing impairment (DFNA4). Am J Hum Genet 74: 770-6

Durocher D, Chen CY, Ardati A, Schwartz RJ, Nemer M (1996) The atrial natriuretic factor promoter is a downstream target for Nkx- 2.5 in the myocardium. Mol Cell Biol 16: 4648-55

Eldadah ZA, Hamosh A, Biery NJ, Montgomery RA, Duke M, Elkins R, Dietz HC (2001) Familial Tetralogy of Fallot caused by mutation in the jagged1 gene. Hum Mol Genet 10: 163-9

Elek C, Vitez M, Czeizel E (1991) [Holt-Oram syndrome]. Orv Hetil 132: 73-4, 77-8.

Ellison RC, Peckham GJ, Lang P, Talner NS, Lerer TJ, Lin L, Dooley KJ, Nadas AS (1983) Evaluation of the preterm infant for patent ductus arteriosus. Pediatrics 71: 364-72

Emanuel R, Somerville J, Inns A, Withers R (1983) Evidence of congenital heart disease in the offspring *of parents with* atrioventricular defects. Br Heart J 49: 144-7

Emmanouilides GC, Linde LM, Crittenden IH (1964) Pulmonary Artery Stenosis Associated with Ductus Arteriosus Following Maternal Rubella. Circulation 29: SUPPL:514-22

Fan C, Duhagon MA, Oberti C, Chen S, Hiroi Y, Komuro I, Duhagon PI, Canessa R, Wang Q (2003) Novel TBX5 mutations and molecular mechanism for Holt-Oram syndrome. J Med Genet 40: e29

Farrance IK, Mar JH, Ordahl CP (1992) M-CAT binding factor is related to the SV40 enhancer binding factor, TEF-1. J Biol Chem 267: 17234-40

Fatkin D, Christe ME, Aristizabal O, McConnell BK, Srinivasan S, Schoen FJ, Seidman CE, Turnbull DH, Seidman JG (1999) Neonatal cardiomyopathy in mice homozygous for the Arg403Gln mutation in the alpha cardiac myosin heavy chain gene. J Clin Invest 103: 147-53

Ferencz C (1993) Epidemiology of congenital heart disease: The Baltimore-Washington infant study 1981-1989. Futura Publishings, Mount Kisco NY

Ferencz C, Neill CA, Boughman JA, Rubin JD, Brenner JI, Perry LW (1989) Congenital cardiovascular malformations associated with chromosome abnormalities: an epidemiologic study. J Pediatr 114: 79-86

Firulli A, McFadden DG, Lin Q, Srivastava D, Olson EN (1998) Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. Nature Genetics 18: 266-270

Fishman MC, Chien KR (1997) Fashioning the vertebrate heart: earliest embryonic decisions. Development 124: 2099-117

Franco D, Campione M, Kelly R, Zammit PS, Buckingham M, Lamers WH, Moorman AF (2000) Multiple transcriptional domains, with distinct left and right components, in the atrial chambers of the developing heart. Circ Res 87: 984-91

Frasch M (1995) Induction of visceral and cardiac mesoderm by ectodermal Dpp in the early Drosophila embryo. Nature 374: 464-7

Funke-Kaiser H, Lemmer J, Langsdorff CV, Thomas A, Kovacevic SD, Strasdat M, Behrouzi T, Zollmann FS, Paul M, Orzechowski HD (2003) Endothelin-converting enzyme-1 (ECE-1) is a downstream target of the homeobox transcription factor Nkx2-5. Faseb J 17: 1487-9

Fyler DC (1980) Report of the New England Regional Infant Cardiac Program. Pediatrics 65: 375-461

Gajewski K, Fossett N, Molkentin JD, Schulz RA (1999) The zinc finger proteins Pannier and GATA4 function as cardiogenic factors in Drosophila. Development 126: 5679-88

Gajewski K, Zhang Q, Choi CY, Fossett N, Dang A, Kim YH, Kim Y, Schulz RA (2001) Pannier is a transcriptional target and partner of Tinman during Drosophila cardiogenesis. Dev Biol 233: 425-36

Galvin KM, Donovan MJ, Lynch CA, Meyer RI, Paul RJ, Lorenz JN, Fairchild-Huntress V, Dixon KL, Dunmore JH, Gimbrone MA, Falb D, Huszar D (2000) A role for Smad6 in development and homeostasis of the cardiovascular system. Nature Genetics 24: 171-174

Garcia-Martinez V, Schoenwolf GC (1993) Primitive-streak origin of the cardiovascular system in avian embryos. Dev Biol 159: 706-19

Garg V, Kathiriya IS, Barnes R, Schluterman MK, King IN, Butler CA, Rothrock CR, Eapen RS, Hirayama-Yamada K, Joo K, Matsuoka R, Cohen JC, Srivastava D (2003) GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature 424: 443-7

Garrity DM, Childs S, Fishman MC (2002) The heartstrings mutation in zebrafish causes heart/fin Tbx5 deficiency syndrome. Development 129: 4635-45

Geisterfer-Lowrance AA, Kass S, Tanigawa G, Vosberg HP, McKenna W, Seidman CE, Seidman JG (1990) A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. Cell 62: 999-1006

Gomez-Skarmeta JL, Diez del Corral R, de la Calle-Mustienes E, Ferre-Marco D, Modolell J (1996) Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. Cell 85: 95-105

Gong W, Gottlieb S, Collins J, Blescia A, Dietz H, Goldmuntz E, McDonald-McGinn DM, Zackai EH, Emanuel BS, Driscoll DA, Budarf ML (2001) Mutation analysis of TBX1 in non-deleted patients with features of DGS/VCFS or isolated cardiovascular defects. J Med Genet 38: E45

Gruenauer-Kloevekorn C, Froster UG (2003) Holt-Oram syndrome: a new mutation in the TBX5 gene in two unrelated families. Ann Genet 46: 19-23

Gupta M, Gupta MP (1997) Cardiac hypertrophy: old concepts, new perspectives. Mol Cell Biochem 176: 273-9

Gupta M, Sueblinvong V, Raman J, Jeevanandam V, Gupta MP (2003) Single-stranded DNA-binding proteins PURalpha and PURbeta bind to a purine-rich negative regulatory element of the alpha-myosin heavy chain gene and control transcriptional and translational regulation of the gene expression. Implications in the repression of alpha-myosin heavy chain during heart failure. J Biol Chem 278: 44935-48

Gupta M, Zak R, Libermann TA, Gupta MP (1998) Tissue-restricted expression of the cardiac alpha-myosin heavy chain gene is controlled by a downstream repressor element containing a palindrome of two ets-binding sites. Mol Cell Biol 18: 7243-58

Haddad F, Bodell PW, Qin AX, Giger JM, Baldwin KM (2003) Role of antisense RNA in coordinating cardiac myosin heavy chain gene switching. J Biol Chem 278: 37132-8

Harvey RP (2002) Patterning the vertebrate heart. Nat Rev Genet 3: 544-56

Heathcote K, Braybrook C, Abushaban L, Guy M, Khetyar ME, Patton MA, Carter ND, Scambler PJ, Syrris P (2005) Common arterial trunk associated with a homeodomain mutation of NKX2.6. Hum Mol Genet 14: 585-93

Heinritz W, Moschik A, Kujat A, Spranger S, Heilbronner H, Demuth S, Bier A, Tihanyi M, Mundlos S, Gruenauer-Kloevekorn C, Froster UG (2005) Identification of new mutations in the TBX5 gene in patients with Holt-Oram syndrome. Heart 91: 383-4

Herranz H, Morata G (2001) The functions of pannier during Drosophila embryogenesis. Development 128: 4837-46

Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA (2006) Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat Genet 38: 82-5

Hiroi Y, Kudoh S, Monzen K, Ikeda Y, Yazaki Y, Nagai R, Komuro I (2001) Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. Nat Genet 28: 276-80

Hodge TP, Cross R, Kendrick-Jones J (1992) Role of the COOH-terminal nonhelical tailpiece in the assembly of a vertebrate nonmuscle myosin rod. J Cell Biol 118: 1085-95

Holmes LB (1965) Congenital Heart Disease and Upper-Extremity Deformities: a Report of Two Families. N Engl J Med 272: 437-44

Holmes LB, Harvey EA, Coull BA, Huntington KB, Khoshbin S, Hayes AM, Ryan LM (2001) The teratogenicity of anticonvulsant drugs. N Engl J Med 344: 1132-8

Holt M, Oram S (1960) Familial heart disease with skeletal malformations. Br Heart J 22: 236-42

Houdusse A, Kalabokis VN, Himmel D, Szent-Gyorgyi AG, Cohen C (1999) Atomic structure of scallop myosin subfragment S1 complexed with MgADP: a novel conformation of the myosin head. Cell 97: 459-70

Icardo JM, Ojeda JL (1984) Effects of colchicine on the formation and looping of the tubular heart of the embryonic chick. Acta Anat (Basel) 119: 1-9

Ichida F, Tsubata S, Bowles KR, Haneda N, Uese K, Miyawaki T, Dreyer WJ, Messina J, Li H, Bowles NE, Towbin JA (2001) Novel gene mutations in patients with left ventricular noncompaction or Barth syndrome. Circulation 103: 1256-63

Imamura S, Matsuoka R, Hiratsuka E, Kimura M, Nakanishi T, Nishikawa T, Furutani Y, Takao A (1991) Adaptational changes of MHC gene expression and isozyme transition in cardiac overloading. Am J Physiol 260: H73-9

Ingwall JS (2002) Is creatine kinase a target for AMP-activated protein kinase in the heart? J Mol Cell Cardiol 34: 1111-20

Inoue K, Lupski JR (2002) Molecular mechanisms for genomic disorders. Annu Rev Genomics Hum Genet 3: 199-242

Isaac A, Sargent MG, Cooke J (1997) Control of vertebrate left-right asymmetry by a snail-related zinc finger gene. Science 275: 1301-4

Itasaki N, Nakamura H, Sumida H, Yasuda M (1991) Actin bundles on the right side in the caudal part of the heart tube play a role in dextro-looping in the embryonic chick heart. Anat Embryol (Berl) 183: 29-39

Itoh K, Adelstein RS (1995) Neuronal cell expression of inserted isoforms of vertebrate nonmuscle myosin heavy chain II-B. J Biol Chem 270: 14533-40

Izumo S, Mahdavi V (1988) Thyroid hormone receptor alpha isoforms generated by alternative splicing differentially activate myosin HC gene transcription. Nature 334: 539-42

Jones WK, Grupp IL, Doetschman T, Grupp G, Osinska H, Hewett TE, Boivin G, Gulick J, Ng WA, Robbins J (1996) Ablation of the murine alpha myosin heavy chain gene leads to dosage effects and functional deficits in the heart. J Clin Invest 98: 1906-17

Kallen B (1987) Search for teratogenic risks with the aid of malformation registries. Teratology 35: 47-52

Kamisago M, Sharma SD, DePalma SR, Solomon S, Sharma P, McDonough B, Smoot L, Mullen MP, Woolf PK, Wigle ED, Seidman JG, Seidman CE (2000) Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. N Engl J Med 343: 1688-96

Karatza AA, Wolfenden JL, Taylor MJ, Wee L, Fisk NM, Gardiner HM (2002) Influence of twin-twin transfusion syndrome on fetal cardiovascular structure and function: prospective case-control study of 136 monochorionic twin pregnancies. Heart 88: 271-7

Kelly RG, Brown NA, Buckingham ME (2001) The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm. Dev Cell 1: 435-40

Khan SG, Metin A, Gozukara E, Inui H, Shahlavi T, Muniz-Medina V, Baker CC, Ueda T, Aiken JR, Schneider TD, Kraemer KH (2004) Two essential splice lariat branchpoint sequences in one intron in a xeroderma pigmentosum DNA repair gene: mutations result in reduced XPC mRNA levels that correlate with cancer risk. Hum Mol Genet 13: 343-52

Khoury MJ, Cordero JF, Greenberg F, James LM, Erickson JD (1983) A population study of the VACTERL association: evidence for its etiologic heterogeneity. Pediatrics 71: 815-20

Kim JS, Viragh S, Moorman AF, Anderson RH, Lamers WH (2001) Development of the myocardium of the atrioventricular canal and the vestibular spine in the human heart. Circ Res 88: 395-402

Kitajima S, Takagi A, Inoue T, Saga Y (2000) MesP1 and MesP2 are essential for the development of cardiac mesoderm. Development 127: 3215-26

Klinedinst SL, Bodmer R (2003) Gata factor Pannier is required to establish competence for heart progenitor formation. Development 130: 3027-38

Knoblauch H, Thiel G, Tinschert S, Korner H, Tennstedt C, Chaoui R, Kohlhase J, Dixkens C, Blanck C (2000) Clinical and molecular cytogenetic studies of a large de novo interstitial deletion 16q11.2-16q21 including the putative transcription factor gene SALL1. J Med Genet 37: 389-92

Kohlhase J (2000) SALL1 mutations in Townes-Brocks syndrome and related disorders. Hum Mutat 16: 460-6

Kohlhase J, Schubert L, Liebers M, Rauch A, Becker K, Mohammed SN, Newbury-Ecob R, Reardon W (2003) Mutations at the SALL4 locus on chromosome 20 result in a range of clinically overlapping phenotypes, including Okihiro syndrome, Holt-Oram syndrome, acro-renal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. J Med Genet 40: 473-8

Kohlhase J, Wischermann A, Reichenbach H, Froster U, Engel W (1998) Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. Nat Genet 18: 81-3

Koivurova S, Hartikainen AL, Gissler M, Hemminki E, Sovio U, Jarvelin MR (2002) Neonatal outcome and congenital malformations in children born after in-vitro fertilization. Hum Reprod 17: 1391-8

Komuro I, Izumo S (1993) Csx: a murine homeobox-containing gene specifically expressed in the developing heart. Proc Natl Acad Sci U S A 90: 8145-9

Korn ED (2000) Coevolution of head, neck, and tail domains of myosin heavy chains. Proc Natl Acad Sci U S A 97: 12559-64

Kosaki R, Gebbia M, Kosaki K, Lewin M, Bowers P, Towbin JA, Casey B (1999) Left-right axis malformations associated with mutations in ACVR2B, the gene for human activin receptor type IIB. Am J Med Genet 82: 70-6

Koyanagi M, Haendeler J, Badorff C, Brandes RP, Hoffmann J, Pandur P, Zeiher AM, Kuhl M, Dimmeler S (2005) Non-canonical Wnt signaling enhances differentiation of human circulating progenitor cells to cardiomyogenic cells. J Biol Chem 280: 16838-42

Kralovicova J, Houngninou-Molango S, Kramer A, Vorechovsky I (2004) Branch site haplotypes that control alternative splicing. Hum Mol Genet 13: 3189-202

Krenz M, Sanbe A, Bouyer-Dalloz F, Gulick J, Klevitsky R, Hewett TE, Osinska HE, Lorenz JN, Brosseau C, Federico A, Alpert NR, Warshaw DM, Perryman MB, Helmke SM, Robbins J (2003) Analysis of myosin heavy chain functionality in the heart. J Biol Chem 278: 17466-74

Kuivenhoven JA, Weibusch H, Pritchard PH, Funke H, Benne R, Assmann G, Kastelein JJ (1996) An intronic mutation in a lariat branchpoint sequence is a direct cause of an inherited human disorder (fish-eye disease). J Clin Invest 98: 358-64

Kupperman E, An S, Osborne N, Waldron S, Stainier DY (2000) A sphingosine-1-phosphate receptor regulates cell migration during vertebrate heart development. Nature 406: 192-5

Kurabayashi M, Tsuchimochi H, Komuro I, Takaku F, Yazaki Y (1988) Molecular cloning and characterization of human cardiac alpha- and beta-form myosin heavy chain complementary DNA clones. Regulation of expression during development and pressure overload in human atrium. J Clin Invest 82: 524-31

Lalwani AK, Goldstein JA, Kelley MJ, Luxford W, Castelein CM, Mhatre AN (2000) Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in nonmuscle myosin MYH9. Am J Hum Genet 67: 1121-8

Lamers WH, Moorman AF (2002) Cardiac septation: a late contribution of the embryonic primary myocardium to heart morphogenesis. Circ Res 91: 93-103

Lamers WH, Wessels A, Verbeek FJ, Moorman AF, Viragh S, Wenink AC, Gittenberger-de Groot AC, Anderson RH (1992) New findings concerning ventricular septation in the human heart. Implications for maldevelopment. Circulation 86: 1194-205

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265: 2037-48

Larsen WJ (2001) Human embryology. Churchill Livingstone, Philadelphia

Le Meur N, Goldenberg A, Michel-Adde C, Drouin-Garraud V, Blaysat G, Marret S, Amara SA, Moirot H, Joly-Helas G, Mace B, Kleinfinger P, Saugier-Veber P, Frebourg T, Rossi A (2005) Molecular characterization of a 14q deletion in a boy with features of Holt-Oram syndrome. Am J Med Genet A 134: 439-42

Lee KF, Simon H, Chen H, Bates B, Hung MC, Hauser C (1995) Requirement for neuregulin receptor erbB2 in neural and cardiac development. Nature 378: 394-8

Lenke RR, Levy HL (1980) Maternal phenylketonuria and hyperphenylalaninemia. An international survey of the outcome of untreated and treated pregnancies. N Engl J Med 303: 1202-8

Levin M (1997) Left-right asymmetry in vertebrate embryogenesis. Bioessays 19: 287-96

Levin M, Johnson RL, Stern CD, Kuehn M, Tabin C (1995) A molecular pathway determining left-right asymmetry in chick embryogenesis. Cell 82: 803-14

Lewis KB, Bruce RA, Baum D, Motulsky AG (1965a) The upper limb-cardiovascular syndrome. An autosomal dominant genetic effect on embryogenesis. Jama 193: 1080-6

Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis AR, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, Bonnet D, Lyonnet S, Young ID, Raeburn JA, Buckler AJ, Law DJ, Brook JD (1997) Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. Nat Genet 15: 21-9

Li S, Zhou D, Lu MM, Morrisey EE (2004) Advanced cardiac morphogenesis does not require heart tube fusion. Science 305: 1619-22

Liberatore CM, Searcy-Schrick RD, Vincent EB, Yutzey KE (2002) Nkx-2.5 gene induction in mice is mediated by a Smad consensus regulatory region. Dev Biol 244: 243-56

Lin AE (1990) Congenital heart defects in malformation syndromes. Clin Perinatol 17: 641-73

Lin Q, Schwarz J, Bucana C, Olson EN (1997) Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. Science 276: 1404-1407

Linask KK, Knudsen KA, Gui YH (1997) N-cadherin-catenin interaction: necessary component of cardiac cell compartmentalization during early vertebrate heart development. Dev Biol 185: 148-64

Linask KK, Lash JW (1986) Precardiac cell migration: fibronectin localization at mesoderm-endoderm interface during directional movement. Dev Biol 114: 87-101

Linask KK, Yu X, Chen Y, Han MD (2002) Directionality of heart looping: effects of Pitx2c misexpression on flectin asymmetry and midline structures. Dev Biol 246: 407-17

Liu XZ, Walsh J, Tamagawa Y, Kitamura K, Nishizawa M, Steel KP, Brown SD (1997) Autosomal dominant non-syndromic deafness caused by a mutation in the myosin VIIA gene. Nat Genet 17: 268-9

Lockwood WK, Bodmer R *(2002) The* patterns of wingless, decapentaplegic, and tinman position the Drosophila heart. Mech *Dev 114: 13-26*

Lompre AM, Nadal-Ginard B, Mahdavi V (1984) Expression of the cardiac ventricular alpha- and beta-myosin heavy chain genes is developmentally and hormonally regulated. J Biol Chem 259: 6437-46

Lowe LA, Yamada S, Kuehn MR (2001) Genetic dissection of nodal function in patterning the mouse embryo. Development 128: 1831-43

Luo ZX, Ji Q, Wible JR, Yuan CX (2003) An Early Cretaceous tribosphenic mammal and metatherian evolution. Science 302: 1934-40

Lyons GE (1996) Vertebrate heart development. Current Opinion in Genetics & Development 6: 454-460

Lyons GE, Schiaffino S, Sassoon D, Barton P, Buckingham M (1990) Developmental regulation of myosin gene expression in mouse cardiac muscle. J Cell Biol 111: 2427-36

Manning A, McLachlan JC (1990) Looping of chick embryo hearts in vitro. J Anat 168: 257-63

Manning N, Archer N (2006) A study to determine the incidence of structural congenital heart disease in monochorionic twins. Prenat Diagn

Margulies EH, Kardia SL, Innis JW (2001) A comparative molecular analysis of developing mouse forelimbs and hindlimbs using serial analysis of gene expression (SAGE). Genome Res 11: 1686-98

Mariner PD, Luckey SW, Long CS, Sucharov CC, Leinwand LA (2005) Yin Yang 1 represses alpha-myosin heavy chain gene expression in pathologic cardiac hypertrophy. Biochem Biophys Res Commun 326: 79-86

Martin J, Han C, Gordon LA, Terry A, Prabhakar S, She X, Xie G, Hellsten U, Chan YM, Altherr M, Couronne O, Aerts A, Bajorek E, Black S, Blumer H, Branscomb E, Brown NC, Bruno WJ, Buckingham JM, Callen DF, Campbell CS, Campbell ML, Campbell EW, Caoile C, Challacombe JF, Chasteen LA, Chertkov O, Chi HC, Christensen M, Clark LM, Cohn JD, Denys M, Detter JC, Dickson M, Dimitrijevic-Bussod M, Escobar J, Fawcett JJ, Flowers D, Fotopulos D, Glavina T, Gomez M, Gonzales E, Goodstein D, Goodwin LA, Grady DL, Grigoriev I, Groza M, Hammon N, Hawkins T, Haydu L, Hildebrand CE, Huang W, Israni S, Jett J, Jewett PB, Kadner K, Kimball H, Kobayashi A, Krawczyk MC, Leyba T, Longmire JL, Lopez F, Lou Y, Lowry S, Ludeman T, Manohar CF, Mark GA, McMurray KL, Meincke LJ, Morgan J, Moyzis RK, Mundt MO, Munk AC, Nandkeshwar RD, Pitluck S, Pollard M, Predki P, Parson-Quintana B, Ramirez L, Rash S, Retterer J, Ricke DO, Robinson DL, Rodriguez A, Salamov A, Saunders EH, Scott D, Shough T, Stallings RL, Stalvey M, Sutherland RD, Tapia R, Tesmer JG, Thayer N, Thompson LS, Tice H, Torney DC, Tran-Gyamfi M, Tsai M, Ulanovsky LE, et al. (2004) The sequence and analysis of duplication-rich human chromosome 16. Nature 432: 988-94

Martinsson T, Oldfors A, Darin N, Berg K, Tajsharghi H, Kyllerman M, Wahlstrom J (2000) Autosomal dominant myopathy: missense mutation (Glu-706 --> Lys) in the myosin heavy chain IIa gene. Proc Natl Acad Sci U S A 97: 14614-9

Marvin MJ, Di Rocco G, Gardiner A, Bush SM, Lassar AB (2001) Inhibition of Wnt activity induces heart formation from posterior mesoderm. Genes Dev 15: 316-27

McDermott DA, Bressan MC, He J, Lee JS, Aftimos S, Brueckner M, Gilbert F, Graham GE, Hannibal MC, Innis JW, Pierpont ME, Raas-Rothschild A, Shanske AL, Smith WE, Spencer RH, St John-Sutton MG, van Maldergem L, Waggoner DJ, Weber M, Basson CT (2005) TBX5 genetic testing validates strict clinical criteria for Holt-Oram syndrome. Pediatr Res 58: 981-6

Meilhac SM, Esner M, Kelly RG, Nicolas JF, Buckingham ME (2004) The clonal origin of myocardial cells in different regions of the embryonic mouse heart. Dev Cell 6: 685-98

Melchionda S, Ahituv N, Bisceglia L, Sobe T, Glaser F, Rabionet R, Arbones ML, Notarangelo A, Di Iorio E, Carella M, Zelante L, Estivill X, Avraham KB, Gasparini P (2001) MYO6, the human homologue of the gene responsible for deafness in Snell's waltzer mice, is mutated in autosomal dominant nonsyndromic hearing loss. Am J Hum Genet 69: 635-40

Mengel MB, Searight HR, Cook K (2006) Preventing Alcohol-exposed Pregnancies. J Am Board Fam Med 19: 494-505

Meredith C, Herrmann R, Parry C, Liyanage K, Dye DE, Durling HJ, Duff RM, Beckman K, de Visser M, van der Graaff MM, Hedera P, Fink JK, Petty EM, Lamont P, Fabian V, Bridges L, Voit T, Mastaglia FL, Laing NG (2004) Mutations in the slow skeletal muscle fiber myosin heavy chain gene (MYH7) cause laing early-onset distal myopathy (MPD1). Am J Hum Genet 75: 703-8

Mermall V, Post PL, Mooseker MS (1998) Unconventional myosins in cell movement, membrane traffic, and signal transduction. Science 279: 527-33

Minehardt TJ, Marzari N, Cooke R, Pate E, Kollman PA, Car R (2002) A classical and ab initio study of the interaction of the myosin triphosphate binding domain with ATP. Biophys J 82: 660-75

Miyata S, Minobe W, Bristow MR, Leinwand LA (2000) Myosin heavy chain isoform expression in the failing and nonfailing human heart. Circ Res 86: 386-90

Mjaatvedt CH, Nakaoka T, Moreno-Rodriguez R, Norris RA, Kern MJ, Eisenberg CA, Turner D, Markwald RR (2001) The outflow tract of the heart is recruited from a novel heart-forming field. Dev Biol 238: 97-109

Molkentin JD, Jobe SM, Markham BE (1996) Alpha-myosin heavy chain gene regulation: delineation and characterization of the cardiac muscle-specific enhancer and muscle-specific promoter. J Mol Cell Cardiol 28: 1211-25

Molkentin JD, Kalvakolanu DV, Markham BE (1994) Transcription factor GATA-4 regulates cardiac muscle-specific expression of the alpha-myosin heavy-chain gene. Mol Cell Biol 14: 4947-57

Molkentin JD, Markham BE (1993) Myocyte-specific enhancer-binding factor (MEF-2) regulates alpha-cardiac myosin heavy chain gene expression in vitro and in vivo. J Biol Chem 268: 19512-20

Molkentin JD, Markham BE (1994) An M-CAT binding factor and an RSRF-related A-rich binding factor positively regulate expression of the alpha-cardiac myosin heavy-chain gene in vivo. Mol Cell Biol 14: 5056-65

Monteiro de Pina-Neto J (1984) Phenotypic variability in Townes-Brocks syndrome. Am J Med Genet 18: 147-52

Monzen K, Hiroi Y, Kudoh S, Akazawa H, Oka T, Takimoto E, Hayashi D, Hosoda T, Kawabata M, Miyazono K, Ishii S, Yazaki Y, Nagai R, Komuro I (2001) Smads, TAK1, and their common target ATF-2 play a critical role in cardiomyocyte differentiation. J Cell Biol 153: 687-98

Monzen K, Shiojima I, Hiroi Y, Kudoh S, Oka T, Takimoto E, Hayashi D, Hosoda T, Habara-Ohkubo A, Nakaoka T, Fujita T, Yazaki Y, Komuro I (1999) Bone morphogenetic proteins induce cardiomyocyte differentiation through the mitogen-activated protein kinase kinase kinase TAK1 and cardiac transcription factors Csx/Nkx-2.5 and GATA-4. Mol Cell Biol 19: 7096-105

Moorman A, Webb S, Brown NA, Lamers W, Anderson RH (2003) Development of the heart: (1) formation of the cardiac chambers and arterial trunks. Heart 89: 806-14

Moorman AF, Schumacher CA, de Boer PA, Hagoort J, Bezstarosti K, van den Hoff MJ, Wagenaar GT, Lamers JM, Wuytack F, Christoffels VM, Fiolet JW (2000) Presence of functional sarcoplasmic reticulum in the developing heart and its confinement to chamber myocardium. Dev Biol 223: 279-90

Mori AD, Bruneau BG (2004) TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol 19: 211-5

Morkin E (2000) Control of cardiac myosin heavy chain gene expression. Microsc Res Tech 50: 522-31

Morkin E, Flink IL, Goldman S (1983) Biochemical and physiologic effects of thyroid hormone on cardiac performance. Prog Cardiovasc Dis 25: 435-64

Mornet D, Bertrand RU, Pantel P, Audemard E, Kassab R (1981) Proteolytic approach to structure and function of actin recognition site in myosin heads. Biochemistry 20: 2110-20

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet 7: 277-318

Moss AJ (1992) Clues in diagnosing congenital heart disease. West J Med 156: 392-8

Mummenhoff J, Houweling AC, Peters T, Christoffels VM, Ruther U (2001) Expression of Irx6 during mouse morphogenesis. Mech Dev 103: 193-5

Muncke N, Jung C, Rudiger H, Ulmer H, Roeth R, Hubert A, Goldmuntz E, Driscoll D, Goodship J, Schon K, Rappold G (2003) Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). Circulation 108: 2843-50

Murphy CT, Spudich JA (1998) Dictyostelium myosin 25-50K loop substitutions specifically affect ADP release rates. Biochemistry 37: 6738-44

Murphy CT, Spudich JA (2000) Variable surface loops and myosin activity: accessories to a motor. J Muscle Res Cell Motil 21: 139-51

Narita N, Bielinska M, Wilson DB (1997) Wild-type endoderm abrogates the ventral developmental defects associated with GATA-4 deficiency in the mouse. Dev Biol 189: 270-4

Navankasattusas S, Sawadogo M, van Bilsen M, Dang CV, Chien KR (1994) The basic helix-loop-helix protein upstream stimulating factor regulates the cardiac ventricular myosin light-chain 2 gene via independent cis regulatory elements. Mol Cell Biol 14: 7331-9

Nelis E, Van Broeckhoven C, De Jonghe P, Lofgren A, Vandenberghe A, Latour P, Le Guern E, Brice A, Mostacciuolo ML, Schiavon F, Palau F, Bort S, Upadhyaya M, Rocchi M, Archidiacono N, Mandich P, Bellone E, Silander K, Savontaus ML, Navon R, Goldberg-Stern H, Estivill X, Volpini V, Friedl W, Gal A, et al. (1996) Estimation of the mutation frequencies in Charcot-Marie-Tooth disease type 1 and hereditary neuropathy with liability to pressure palsies: a European collaborative study. Eur J Hum Genet 4: 25-33

Newbury-Ecob RA, Leanage R, Raeburn JA, Young ID (1996) Holt-Oram syndrome: a clinical genetic study. J Med Genet 33: 300-7

Niederreither K, Ward SJ, Dolle P, Chambon P (1996) Morphological and molecular characterization of retinoic acid- induced limb duplications in mice. Developmental Biology 176: 185-198

Niimura H, Patton KK, McKenna WJ, Soults J, Maron BJ, Seidman JG, Seidman CE (2002) Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly. Circulation 105: 446-51

Nonaka S, Shiratori H, Saijoh Y, Hamada H (2002) Determination of left-right patterning of the mouse embryo by artificial nodal flow. Nature 418: 96-9

Nonaka S, Tanaka Y, Okada Y, Takeda S, Harada A, Kanai Y, Kido M, Hirokawa N (1998) Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell 95: 829-837

Nonaka S, Yoshiba S, Watanabe D, Ikeuchi S, Goto T, Marshall WF, Hamada H (2005) De novo formation of left-right asymmetry by posterior tilt of nodal cilia. PLoS Biol 3: e268

Nora JJ (1993) Causes of congenital heart diseases: old and new modes, mechanisms, and models. Am Heart J 125: 1409-19

Ockey CH, Feldman GV, Macaulay ME, Delaney MJ (1967) A large deletion of the long arm of chromosome No. 4 in a child with limb abnormalities. Arch Dis Child 42: 428-34

Ogura K, Matsumoto K, Kuroiwa A, Isobe T, Otoguro T, Jurecic V, Baldini A, Matsuda Y, Ogura T (2001) Cloning and chromosome mapping of human and chicken Iroquois (IRX) genes. Cytogenet Cell Genet 92: 320-5

Ott J (1976) A computer program for linkage analysis of general human pedigrees. Am J Hum Genet 28: 528-9

Palmiter KA, Tyska MJ, Dupuis DE, Alpert NR, Warshaw DM (1999) Kinetic differences at the single molecule level account for the functional diversity of rabbit cardiac myosin isoforms. J Physiol 519 Pt 3: 669-78

Park M, Wu X, Golden K, Axelrod JD, Bodmer R (1996) The wingless signaling pathway is directly involved in Drosophila heart development. Dev Biol 177: 104-16

Pastural E, Barrat FJ, Dufourcq-Lagelouse R, Certain S, Sanal O, Jabado N, Seger R, Griscelli C, Fischer A, de Saint Basile G (1997) Griscelli disease maps to chromosome 15q21 and is associated with mutations in the myosin-Va gene. Nat Genet 16: 289-92

Penaloza D, Arias-Stella J, Sime F, Recavarren S, Marticorena E (1964) The Heart and Pulmonary Circulation in Children at High Altitudes: Physiological, Anatomical, and Clinical Observations. Pediatrics 34: 568-82

Peters T, Ausmeier K, Dildrop R, Ruther U (2002) The mouse Fused toes (Ft) mutation is the result of a 1.6-Mb deletion including the entire Iroquois B gene cluster. Mamm Genome 13: 186-8

Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA (2002) Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31: 33-6

Pizzuti A, Sarkozy A, Newton AL, Conti E, Flex E, Digilio MC, Amati F, Gianni D, Tandoi C, Marino B, Crossley M, Dallapiccola B (2003) Mutations of ZFPM2/FOG2 gene in sporadic cases of tetralogy of Fallot. Hum Mutat 22: 372-7

Plageman TF, Jr., Yutzey KE (2004) Differential expression and function of Tbx5 and Tbx20 in cardiac development. J Biol Chem 279: 19026-34

Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Sellers JR, Fananapazir L, Epstein ND (1996) Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. Nat Genet 13: 63-9

Ponticos M, Partridge T, Black CM, Abraham DJ, Bou-Gharios G (2004) Regulation of collagen type I in vascular smooth muscle cells by competition between Nkx2.5 and deltaEF1/ZEB1. Mol Cell Biol 24: 6151-61

Poznanski AK, Gall JC, Jr., Stern AM (1970) Skeletal manifestations of the Holt-Oram syndrome. Radiology 94: 45-53.

Putnam EA, Park ES, Aalfs CM, Hennekam RC, Milewicz DM (1997) Parental somatic and germ-line mosaicism for a FBN2 mutation and analysis of FBN2 transcript levels in dermal fibroblasts. Am J Hum Genet 60: 818-27

Ranger AM, Grusby MJ, Hodge MR, Gravallese EM, de la Brousse FC, Hoey T, Mickanin C, Baldwin HS, Glimcher LH (1998) The transcription factor NF-ATc is essential for cardiac valve formation. Nature 392: 186-190

Rayment I, Holden HM, Whittaker M, Yohn CB, Lorenz M, Holmes KC, Milligan RA (1993a) Structure of the actin-myosin complex and its implications for muscle contraction. Science 261: 58-65

Rayment I, Rypniewski WR, Schmidt-Base K, Smith R, Tomchick DR, Benning MM, Winkelmann DA, Wesenberg G, Holden HM (1993b) Three-dimensional structure of myosin subfragment-1: a molecular motor. Science 261: 50-8

Reamon-Buettner SM, Borlak J (2004) TBX5 mutations in non-Holt-Oram syndrome (HOS) malformed hearts. Hum Mutat 24: 104

Reamon-Buettner SM, Borlak J (2006) HEY2 mutations in malformed hearts. Hum Mutat 27: 118

Redkar A, Montgomery M, Litvin J (2001) Fate map of early avian cardiac progenitor cells. Development 128: 2269-79

Reid IS, Turner G (1976) Familial and abnormality. J Pediatr 88: 992-4

Reiter JF, Alexander J, Rodaway A, Yelon D, Patient R, Holder N, Stainier DY (1999) Gata5 is required for the development of the heart and endoderm in zebrafish. Genes Dev 13: 2983-95

Riazi AM, Lee H, Hsu C, Van Arsdell G (2005) CSX/Nkx2.5 modulates differentiation of skeletal myoblasts and promotes differentiation into neuronal cells in vitro. J Biol Chem 280: 10716-20

Ritchie ME (1996) Characterization of human B creatine kinase gene regulation in the heart in vitro and in vivo. J Biol Chem 271: 25485-91

Roberts KE, McElroy JJ, Wong WP, Yen E, Widlitz A, Barst RJ, Knowles JA, Morse JH (2004) BMPR2 mutations in pulmonary arterial hypertension with congenital heart disease. Eur Respir J 24: 371-4

Robinson SW, Morris CD, Goldmuntz E, Reller MD, Jones MA, Steiner RD, Maslen CL (2003) Missense mutations in CRELD1 are associated with cardiac atrioventricular septal defects. Am J Hum Genet 72: 1047-52

Roebroek AJ, Umans L, Pauli IG, Robertson EJ, van Leuven F, Van de Ven WJ, Constam DB (1998) Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase Furin. Development 125: 4863-76

Rones MS, McLaughlin KA, Raffin M, Mercola M (2000) Serrate and Notch specify cell fates in the heart field by suppressing cardiomyogenesis. Development 127: 3865-76

Rosenquist GC (1970) Location and movements of cardiogenic cells in the chick embryo: the heart-forming portion of the primitive streak. Dev Biol 22: 461-75

Rowe RD (1963) Maternal Rubella and Pulmonary Artery Stenoses. Report of Eleven Cases. Pediatrics 32: 180-5

Rowland TW, Hubbell JP, Jr., Nadas AS (1973) Congenital heart disease in infants of diabetic mothers. J Pediatr 83: 815-20

Rowold DJ, Herrera RJ (2000) Alu elements and the human genome. Genetica 108: 57-72

Rozen S, Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers In: Krawetz S. MS (ed) *Bioinformatics Methods and Protocols: Methods in Molecular Biology* Humana Press, Totowa NJ

Rybak M, Kozlowski K, Kleczkowska A, Lewandowska J, Sokolowski J, Soltysik-Wilk E (1971) Holt-Oram syndrome associated with ectromelia and chromosomal aberrations. Am J Dis Child 121: 490-5.

Saga Y, Miyagawa-Tomita S, Takagi A, Kitajima S, Miyazaki J, Inoue T (1999) MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. Development 126: 3437-47

Sambrook J (2001) Molecular Cloning, III edn. Cold Spring Harbor Laboratory Press

Schlange T, Andree B, Arnold H, Brand T (2000a) Expression analysis of the chicken homologue of CITED2 during early stages of embryonic development. Mech Dev 98: 157-60

Schlange T, Andree B, Arnold HH, Brand T (2000b) BMP2 is required for early heart development during a distinct time period. Mech Dev 91: 259-70

Schneider VA, Mercola M (2001) Wnt antagonism initiates cardiogenesis in Xenopus laevis. Genes Dev 15: 304-15

Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, Maron BJ, Seidman CE, Seidman JG (1998) Congenital heart disease caused by mutations in the transcription factor NKX2-5 [see comments]. Science 281: 108-11

Schultheiss TM, Burch JB, Lassar AB (1997) A role for bone morphogenetic proteins in the induction of cardiac myogenesis. Genes Dev 11: 451-62

Schultheiss TM, Lassar, A. B. (1999) Vertebrate heart induction. In: Harvey RP, Rosenthal, N. (ed) In: Heart development. Academic Press, San Diego

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M (2004) Large-scale copy number polymorphism in the human genome. Science 305: 525-8

Sellers JR (2000) Myosins: a diverse superfamily. Biochim Biophys Acta 1496: 3-22

Sen SK, Han K, Wang J, Lee J, Wang H, Callinan PA, Dyer M, Cordaux R, Liang P, Batzer MA (2006) Human genomic deletions mediated by recombination between Alu elements. Am J Hum Genet 79: 41-53

Seri M, Cusano R, Gangarossa S, Caridi G, Bordo D, Lo Nigro C, Ghiggeri GM, Ravazzolo R, Savino M, Del Vecchio M, d'Apolito M, Iolascon A, Zelante

LL, Savoia A, Balduini CL, Noris P, Magrini U, Belletti S, Heath KE, Babcock M, Glucksman MJ, Aliprandis E, Bizzaro N, Desnick RJ, Martignetti JA (2000) Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and Sebastian syndromes. The May-Hegglin/Fechtner Syndrome Consortium. Nat Genet 26: 103-5

Silver W, Steier M, Schwartz O, Zeichner MB (1972) The Holt-Oram syndrome with previously undescribed associated anomalies. Am J Dis Child 124: 911-4.

Simoes-Costa MS, Vasconcelos M, Sampaio AC, Cravo RM, Linhares VL, Hochgreb T, Yan CY, Davidson B, Xavier-Neto J (2005) The evolutionary origin of cardiac chambers. Dev Biol 277: 1-15

Smith AT, Sack GH, Jr., Taylor GJ (1979) Holt-Oram syndrome. J Pediatr 95: 538-43.

Sperling S, Grimm CH, Dunkel I, Mebus S, Sperling HP, Ebner A, Galli R, Lehrach H, Fusch C, Berger F, Hammer S (2005) Identification and functional analysis of CITED2 mutations in patients with congenital heart defects. Hum Mutat 26: 575-82

Splitt M, Wright C, Sen D, Goodship J (1999) Left-isomerism sequence and maternal type-1 diabetes. Lancet 354: 305-6

Srivastava D, Cserjesi P, Olson EN (1995) A Subclass of Bhlh Proteins Required for Cardiac Morphogenesis. Science 270: 1995-1999

Srivastava D, Olson EN (2000) A genetic blueprint for cardiac development. Nature 407: 221-226

Srivastava D, Thomas T, Lin Q, Kirby ML, Brown D, Olson EN (1997) Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. Nature Genetics 16: 154-160

Stalsberg H (1969) The origin of heart asymmetry: right and left contributions to the early chick embryo heart. Dev Biol 19: 109-27

Stalsberg H (1970) Development and ultrastructure of the embryonic heart. II. Mechanism of dextral looping of the embryonic heart. Am J Cardiol 25: 265-71

Sucharov CC, Helmke SM, Langer SJ, Perryman MB, Bristow M, Leinwand L (2004) The Ku protein complex interacts with YY1, is up-regulated in human heart failure, and represses alpha myosin heavy-chain gene expression. Mol Cell Biol 24: 8705-15

Sun G, Lewis LE, Huang X, Nguyen Q, Price C, Huang T (2004) TBX5, a gene mutated in Holt-Oram syndrome, is regulated through a GC box and T-box binding elements (TBEs). J Cell Biochem 92: 189-99

Sun X, Meyers EN, Lewandoski M, Martin GR (1999) Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev 13: 1834-46

Suri C, Jones PF, Patan S, Bartunkova S, Maisonpierre PC, Davis S, Sato TN, Yancopoulos GD (1996) Requisite role of Angiopoietin-1, a ligand for the TIE2 receptor, during embryonic angiogenesis. Cell 87: 1171-1180

Surka WS, Kohlhase J, Neunert CE, Schneider DS, Proud VK (2001) Unique family with Townes-Brocks syndrome, SALL1 mutation, and cardiac defects. Am J Med Genet 102: 250-7

Swynghedauw B (1986) Developmental and functional adaptation of contractile proteins in cardiac and skeletal muscles. Physiol Rev 66: 710-71

Szeto DP, Griffin KJ, Kimelman D (2002) HrT is required for cardiovascular development in zebrafish. Development 129: 5093-101

Tajsharghi H, Thornell LE, Lindberg C, Lindvall B, Henriksson KG, Oldfors A (2003) Myosin storage myopathy associated with a heterozygous missense mutation in MYH7. Ann Neurol 54: 494-500

Takeichi M (1991) Cadherin cell adhesion receptors as a morphogenetic regulator. Science 251: 1451-5

Tam PP, Parameswaran M, Kinder SJ, Weinberger RP (1997) The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development 124: 1631-42

Tanaka M, Chen Z, Bartunkova S, Yamasaki N, Izumo S (1999) The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. Development 126: 1269-80

Tanigawa G, Jarcho JA, Kass S, Solomon SD, Vosberg HP, Seidman JG, Seidman CE (1990) A molecular basis for familial hypertrophic cardiomyopathy: an alpha/beta cardiac myosin heavy chain hybrid gene. Cell 62: 991-8

Terrett JA, Newbury-Ecob R, Cross GS, Fenton I, Raeburn JA, Young ID, Brook JD (1994) Holt-Oram syndrome is a genetically heterogeneous disease with one locus mapping to human chromosome 12q. Nat Genet 6: 401-4.

Terrett JA, Newbury-Ecob R, Smith NM, Li QY, Garrett C, Cox P, Bonnet D, Lyonnet S, Munnich A, Buckler AJ, Brook JD (1996) A translocation at 12q2 refines the interval containing the Holt-Oram syndrome 1 gene. Am J Hum Genet 59: 1337-41.

Tezenas Du Montcel S, Mendizabai H, Ayme S, Levy A, Philip N (1996) Prevalence of 22q11 microdeletion. J Med Genet 33: 719

Townes PL, Brocks ER (1972) Hereditary syndrome of imperforate anus with hand, foot, and ear anomalies. J Pediatr 81: 321-6

Turleau C, de Grouchy J, Chavin-Colin F, Dore F, Seger J, Dautzenberg MD, Arthuis M, Jeanson C (1984) Two patients with interstitial del (14q), one with features of Holt- Oram syndrome. Exclusion mapping of PI (alpha-1-antitrypsin). Ann Genet 27: 237-40

Ueyama T, Kasahara H, Ishiwata T, Yamasaki N, Izumo S (2003) Csm, a cardiac-specific isoform of the RNA helicase Mov10l1, is regulated by Nkx2.5 in embryonic heart. J Biol Chem 278: 28750-7

van der Hoeven F, Schimmang T, Volkmann A, Mattei MG, Kyewski B, Ruther U (1994) Programmed cell death is affected in the novel mouse mutant Fused toes (Ft). Development 120: 2601-7

Vatta M, Mohapatra B, Jimenez S, Sanchez X, Faulkner G, Perles Z, Sinagra G, Lin JH, Vu TM, Zhou Q, Bowles KR, Di Lenarda A, Schimmenti L, Fox M, Chrisco MA, Murphy RT, McKenna W, Elliott P, Bowles NE, Chen J, Valle G, Towbin JA (2003) Mutations in Cypher/ZASP in patients with dilated cardiomyopathy and left ventricular non-compaction. J Am Coll Cardiol 42: 2014-27

Veugelers M, Bressan M, McDermott DA, Weremowicz S, Morton CC, Mabry CC, Lefaivre JF, Zunamon A, Destree A, Chaudron JM, Basson CT (2004) Mutation of perinatal myosin heavy chain associated with a Carney complex variant. N Engl J Med 351: 460-9

von Both I, Silvestri C, Erdemir T, Lickert H, Walls JR, Henkelman RM, Rossant J, Harvey RP, Attisano L, Wrana JL (2004) Foxh1 is essential for development of the anterior heart field. Dev Cell 7: 331-45

Vracar-Grabar M, Russell B (2004) Creatine kinase is an alpha myosin heavy chain 3'UTR mRNA binding protein. J Muscle Res Cell Motil 25: 397-404

Waldo KL, Kumiski DH, Wallis KT, Stadt HA, Hutson MR, Platt DH, Kirby ML (2001) Conotruncal myocardium arises from a secondary heart field. Development 128: 3179-88

Walpole IR, Hockey A (1982) Syndrome of imperforate anus, abnormalities of hands and feet, satyr ears, and sensorineural deafness. J Pediatr 100: 250-2

Walsh T, Walsh V, Vreugde S, Hertzano R, Shahin H, Haika S, Lee MK, Kanaan M, King MC, Avraham KB (2002) From flies' eyes to our ears: mutations in a human class III myosin cause progressive nonsyndromic hearing loss DFNB30. Proc Natl Acad Sci U S A 99: 7518-23

Walter JH (1995) Late effects of phenylketonuria. Arch Dis Child 73: 485-6

Wang A, Liang Y, Fridell RA, Probst FJ, Wilcox ER, Touchman JW, Morton CC, Morell RJ, Noben-Trauth K, Camper SA, Friedman TB (1998) Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. Science 280: 1447-51

Ware SM, Peng J, Zhu L, Fernbach S, Colicos S, Casey B, Towbin J, Belmont JW (2004) Identification and functional analysis of ZIC3 mutations in heterotaxy and related congenital heart defects. Am J Hum Genet 74: 93-105

Watkins H, Rosenzweig A, Hwang DS, Levi T, McKenna W, Seidman CE, Seidman JG (1992) Characteristics and prognostic implications of myosin missense mutations in familial hypertrophic cardiomyopathy. N Engl J Med 326: 1108-14

Webb S, Brown NA, Anderson RH (1998) Formation of the atrioventricular septal structures in the normal mouse. Circ Res 82: 645-56

Webb S, Kanani M, Anderson RH, Richardson MK, Brown NA (2001) Development of the human pulmonary vein and its incorporation in the morphologically left atrium. Cardiol Young 11: 632-42

Weil D, Blanchard S, Kaplan J, Guilford P, Gibson F, Walsh J, Mburu P, Varela A, Levilliers J, Weston MD, et al. (1995) Defective myosin VIIA gene responsible for Usher syndrome type 1B. Nature 374: 60-1

Weinstein MR, Goldfield M (1975) Cardiovascular malformations with lithium use during pregnancy. Am J Psychiatry 132: 529-31

Wessels A, Anderson RH, Markwald RR, Webb S, Brown NA, Viragh S, Moorman AFM, Lamers WH (2000) Atrial development in the human heart: An immunohistochemical study with emphasis on the role of mesenchymal tissues. Anatomical Record 259: 288-300

Whittemore R, Hobbins JC, Engle MA (1982) Pregnancy and its outcome in women with and without surgical treatment of congenital heart disease. Am J Cardiol 50: 641-51

Xu H, Morishima M, Wylie JN, Schwartz RJ, Bruneau BG, Lindsay EA, Baldini A (2004) Tbx1 has a dual role in the morphogenesis of the cardiac outflow tract. Development 131: 3217-27

Yang J, Hu D, Xia J, Yang Y, Ying B, Hu J, Zhou X (2000) Three novel TBX5 mutations in Chinese patients with Holt-Oram syndrome. Am J Med Genet 92: 237-40

Yang SP, Sherman S, Derstine JB, Schonberg SA (1990) Holt-Oram Syndrome Gene May Be on Chromosome-20. Pediatric Research 27: A137-A137

Yatskievych TA, Ladd AN, Antin PB (1997) Induction of cardiac myogenesis in avian pregastrula epiblast: the role of the hypoblast and activin. Development 124: 2561-70

Yelon D (2001) Cardiac patterning and morphogenesis in zebrafish. Dev Dyn 222: 552-63

Yelon D, Ticho B, Halpern ME, Ruvinsky I, Ho RK, Silver LM, Stainier DYR (2000) The bHLH transcription factor Hand2 plays parallel roles in zebrafish heart and pectoral fin development. Development 127: 2573-2582

Yin Z, Xu XL, Frasch M (1997) Regulation of the twist target gene tinman by modular cis-regulatory elements during early mesoderm development. Development 124: 4971-82

Zaffran S, Kelly RG, Meilhac SM, Buckingham ME, Brown NA (2004) Right ventricular myocardium derives from the anterior heart field. Circ Res 95: 261-8

Zeterqvist P (1963) The syndrome of familial atrial septal defect, heart arrhythmia and hand malformation (Holt-Oram) in mother and son. Acta Paediatr 52: 115-22

Zhang XM, Ramalho-Santos M, McMahon AP (2001) Smoothened mutants reveal redundant roles for Shh and Ihh signaling including regulation of L/R asymmetry by the mouse node. Cell 105: 781-92

Zou Y, Evans S, Chen J, Kuo HC, Harvey RP, Chien KR (1997) CARP, a cardiac ankyrin repeat protein, is downstream in the Nkx2-5 homeobox gene pathway. Development 124: 793-804