University of Southampton

Division of Human Genetics

Faculty of Medicine, Health and Life Sciences

School of Medicine

Investigation of a critical chromosome 1q congenital heart disease region

by

David R.W. Fowler

A thesis submitted for the degree of Doctor of Philosophy

October 2004

UNIVERSITY OF SOUTHAMPTON

Abstract

FACULTY OF MEDICINE, NURSING AND HEALTH SCIENCES
SCHOOL OF MEDICINE

Doctor of Philosophy

INVESTIGATION OF A CRITICAL CHROMOSOME 1Q CONGENITAL HEART DISEASE REGION

by

David Richard William Fowler

Chromosomal rearrangements with in 1q42-44 give rise to similar patterns of developmental abnormalities. Cardiac defects are a common feature and are the predominant cause of death. Such cases argue for the presence of a crucial cardiac developmental disease gene with in this region; haploinsufficiency of which causes aberrant cardiac development. A postional candidate strategy was adopted to identify disease-susceptibility genes. There is currently little molecular data with in the literature to allow the delineation of a critical CHD region within 1q42-44 for targeting candidate genes. To address this, molecular analysis was carried out on eleven patients with known 1q deletions. The results obtained suggested that there may be two CHD critical regions witin 1q42-44. To refine the critical regions, CHD patients with normal karyotypes were investigated for submicroscopic deletions using loss-of-heterozygosity analayis. None were found. It became apparent that the introduction of a high-throughput, high-resolution technology would facilitate the identification of submicroscopic deletions with in 1q42-44. The novel technique of microarray-comparative genomic hybridisation was therefore established and validated with a known 1q deletion patient. Genes mapping to the 1q disease intervals were assessed for CHD candidacy. The basement membrane protein, nidogen-1, was identified as a strong candidate gene. Mutation screening of nidogen-1 was undertaken in CHD patients. Direct sequencing identified 10 nucleotide changes. None of these were thought to be pathogenic.

Table of contents

Table of figures

Table of tables

Acknowledgements

First and foremost I would like to thank my supervisor, Professor David Wilson, for firstly giving me the opportunity to study for a PhD and for his continual support throughout. Thankfully our senses of humour were at the same level. I wish to thank all members of the Developmental Genetics group at Southampton for their help. I very much enjoyed working with them all. In particular, my thanks to Glenn Renforth and Mirella Spalluto, both of whom developed the tiling BAC contig of the 11q HLH region, for their guidance and advice with techniques that I have used. With regard to the development of the microarray technology, the work would not have been possible without help from our collaborators at the Wellcome Trust Sanger Institute, Nigel Carter, Heike Fiegler and Phillipa Carr, as well as from Kelly Wilkinson and Ros Ganderton in the Human Genetics Division at Southampton. In addition my thanks go to Carolyn Wallis, the divisional secretary, who was always extremely helpful with using the computers.

On a personal note, considerable gratitude must go to my family for their support throughout my PhD.

Abbreviations

AHF	Anterior heart forming field
A/P	Anterior/posterior
APS	Ammonium persulphate
ASD	Atrial septal defect
AVC	Atrioventricular canal
AVSD	Atrioventricular septal defect
BAC	Bacterial artificial chromosome
BLAST	Basic Local Alignment Search Tool
BMP	Bone morphogenic protein
bp	Base pair
cDNA	Complimentary DNA
CHD	Congenital heart disease
CGH	Comparative genomic hybridisation
DAB	Diaminobenezidine
DAPI	Diamidino-2-phenylindole
ddNTP	Dideoxynucleotide triphosphate
dH_2O	Distilled water
dHPLC	Denaturing high performance liquid chromatography
dpc	Days post conception
DGS	DiGeorge syndrome
DEPC	Diethyl pyrocarbonate
DOP	Degenerate oligonucleotide primer
DORV	Double outlet right ventricle
DNA	Deoxyribonucleic acid

dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EBV	Epstein-Barr virus
ECM	Extracelluar matrix
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-mesenchymal transformation
EST	Expressed sequence tag
FGF	Fibroblast growth factor
FISH	Fluorescent *in situ* hybridisation
FITC	Fluorescein isothiocyanate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
H+E	Heamatoxylin and eosin stained
HLHS	Hypoplastic left heart syndrome
HRP	Horse radish peroxidase
JAM	Junction adhesion molecule
LB	Luria Bertani
LOH	Loss of heterozygosity
Mb	Megabase
mRNA	Messenger ribonucleic acid
NCBI	National centre for Biotechnology Information
OFT	Outflow tract
ORF	Open reading frame
PAC	P1 artificial chromosome
PBS	Phosphate buffered saline

PCR	Polymerase chain reaction
PFA	Paraformaldehyde
RT-PCR	Reverse transcriptase – polymerase chain reaction
SMAD	Signalling against decapentalplegic
STS	Sequence tagged site
SSC	Sodium citrate and chloride
SVAS	Supravavular aortic stenosis
TEAA	Triethylammonium acetate
TEMED	Tetramethylethyenediamine
TGF	Transforming growth factor
TISH	Tissue *in situ* hybridisation
TOF	Tetralogy of Fallot
tRNA	Transfer ribonucleic acid
UDB	Unified database
UTR	Untranslated region
VSD	Ventricular septal defect
YAC	Yeast Artificial Chromosome

Chapter 1: Introduction

1.1 Congenital heart disease

The heart is the first organ to form. Cardiac development is a highly intricate and complex process. This is reflected in the high incidence of cardiac malformations resulting from aberrant development. Congenital heart disease (CHD) is the clinical manifestation of abnormalities in cardiac development. It is a collective term for a heterogenous series of malformations affecting the heart which are present at birth. CHD is restricted to structural defects. Congenital arrhythmias such as the long QT syndrome are usually excluded. A definition was proposed by Mitchell describing the condition as 'a gross structural abnormality of the heart or intrathoracic great vessels that is actually or potentially of functional significance' (Mitchell et al., 1971). Clinical severity is varied, ranging from very mild, which may go undetected well into adult life, to extremely severe possibly resulting in miscarriage or neonatal death. CHD is a significant cause of paediatric mortality and morbidity. These abnormalities are the predominant cause of death within the first week of life. Although reports vary, between 8 and 10 in every 100 live-births will have some form of cardiac defect, while 5-10% of prenatal fatalities are attributable to CHD (Hoffman and Kaplan, 2002).

A genetic aetiology to these malformations has been proposed and the recent identification of mutations in developmental genes of affected individuals has confirmed this hypothesis. In the majority of cases the aetiology is unknown. Significant effort has been invested to identify causative genes. However few to date have been found. Technical advances in the field of molecular biology have had a significant impact on this process through permitting a candidate gene approach to identify congenital heart disease genes. Such studies have been facilitated by our growing knowledge of the genes governing cardiogeneis. Improved understanding of how these genes function has given a greater insight into the mechanisms of cardiac development, and how these can go wrong, and has also aided the identification of CHD candidate genes for investigation.

1.2 The human heart

1.2.1 Cardiac anatomy

The heart is an asymmetric structure, which has a polarity defined by three axes: cranial-caudal, dorsal-ventral, and left-right (see fig 1.1). The heart is located within the pericardial cavity, a fluid-filled fibrous sac, which is enclosed by the pericardium. Normal human cardiogenesis leads to the formation of a four chambered heart with two receiving chambers, the atria, and two muscular pumping chambers, the ventricles (see fig 1.2). The left and right atria are separated by a partitioning wall, the interatrial septum, as are the ventricles by the interventricular septum. The atria are separated from the ventricles by the atrioventricular valves. The valves are comprised of flaps, or cusps, of tissue. The free end of each cusp is linked to the ventricular wall via the chordae tendinae. The right atrioventricular valve, the tricuspid valve, is composed of three cusps. The left atrioventricular valve is comprised of two and is referred to as the mitral valve. Their function is to prevent the regurgitation of blood into the atria during ventricular contraction. The large systemic vein, the vena cava, enters into the right atria; the pulmonary vein drains into the left. The pulmonary artery and the large systemic artery, the aorta, join to the right and left ventricle respectively. Further sets of valves, the semi-lunar valves, are found within the arteries. These arterial valves of the outflow tract consist of half moon shaped flaps, which give them their name, and function to prevent backflow of blood into the ventricle. This parallel chamber arrangement allows for the complete separation of the systemic circulation from the pulmonary circulation.

The walls of the atria and ventricles are made up of three distinct layers: the epicardium, the myocardium and the endocardium. The majority of the wall is comprised of the myocardium, a thick, muscular layer made up of specialised muscle cells, the cardiomyocytes. The interior layer is the endocardium, a single-celled delicate sheet of epithelial tissue that is continuous with the endothelium of the vessels entering and leaving the heart. Surrounding the heart is the epicardium. The free surface of the epicardium consists of a single layer of flattened epithelial cells, the mesothelium.

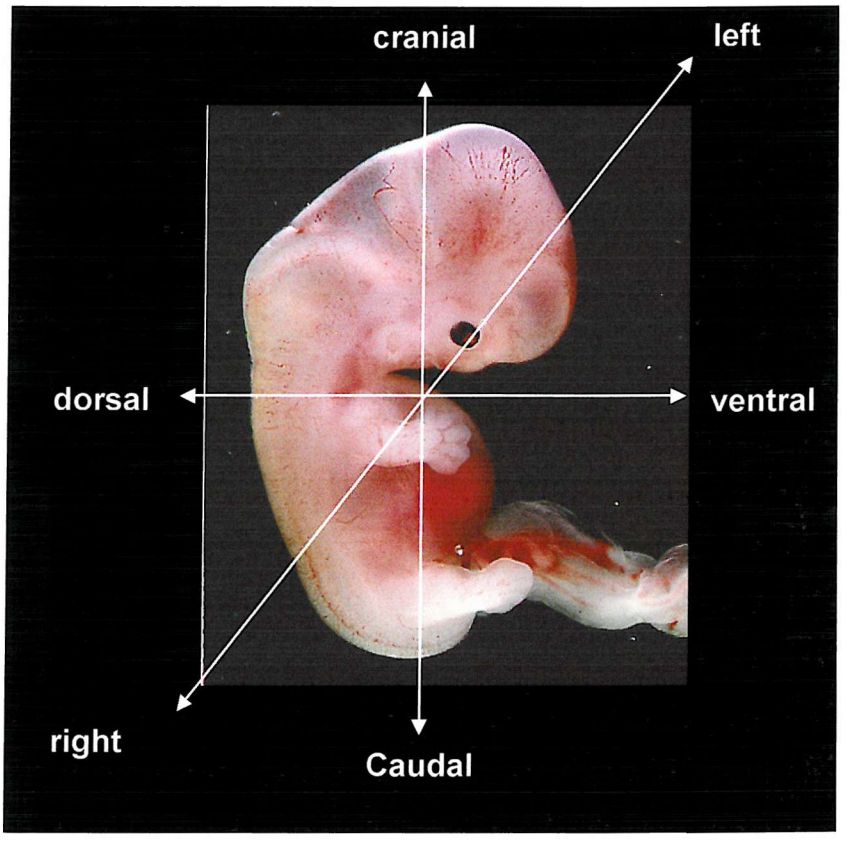

Fig 1.1: A picture indicating the main body axes of a human fetus. The fetus has a polarity defined by three main body axes: cranial to caudal; left to right; dorsal to ventral.

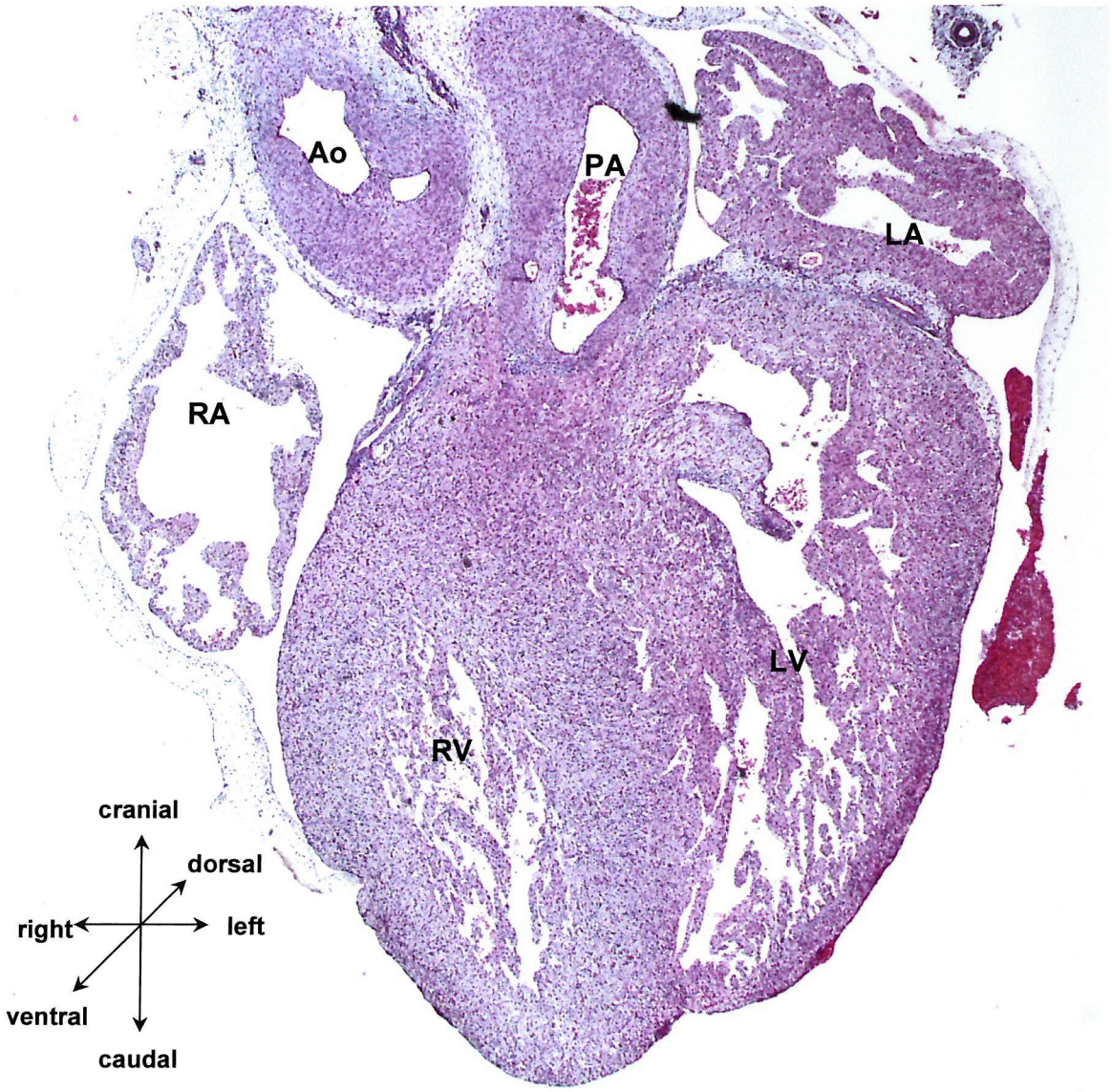

Fig 1.2: A heamatoxylin and eosin (H+E) stained section of an 8 week human fetal heart. LV – left ventricle; **RV –** right ventricle; **LA –** left atria; **RA –** right atria; **Ao –** aorta; **PA –** pulmonary artery.

The mesothelial cells secrete the pericardial fluid which fills the surrounding pericardial cavity to lubricate the movement of the epicardium on the pericardial wall. The walls of the ventricles are comprised of numerous ingrowths, the trabeculae, which aid the rapid distribution of electrical impulses throughout the ventricle.

1.2.2 Cardiac circulation

In the adult heart, deoxygenated blood from the systemic circulation returns to the right atria via the vena cava, from where it passes into the right ventricle. The deoxygenated blood is then pumped to the lungs through the pulmonary artery, where it is re-oxygenated, before returning to the left atria. From here it is pumped into the left ventricle, and re-enters the systemic circulation via the aorta.

The fetal circulation has been adapted to by-pass the lungs, thereby allowing the blood to re-oxygenate at the placenta. Oxygenated blood returns to the fetus from the placenta via the umbilical vein. Blood enters the vena cava, via a vessel called the ductus venosus and passes into the right atrium. A high level of resistance in the pulmonary circulation prevents the majority of blood from entering the right ventricle. Therefore, in contrast to the adult heart, the bulk of the blood is shunted into the left atrium through an opening in the atrial septum called the foramen ovalve. From here oxygenated blood flows in to the left ventricle and in to the systemic circulation via the aorta. Any blood that does enter the pulmonary artery is redistributed to the systemic circulation through the ductus arteriosus, a vessel which connects the pulmonary artery and the aorta. Deoxygenated blood passes in to the umbilical arteries and is re-oxygenated at the placenta. At birth several changes take place to give rise to a separate pulmonary and systemic circulation necessary for an independent existence. The initial inflation of the lungs reduces the pulmonary resistance. As a consequence blood is no longer shunted from the right to the left atria, and passes into the pulmonary circulation via the right ventricle. Oxygenated blood now returns via a functioning pulmonary vein into the left atria resulting in a consequent increase in left atrial pressure forcing the foramen ovalve to close. The ductus arteriosus, which linked the pulmonary artery to the aorta, also closes and disintegrates shortly after birth.

The vessels that make up the circulatory system have a common basic structure. An inner lining comprising a single layer of flattened epithelial cells called the endothelium, supported by a basement membrane and delicate collagenous tissue; this constitutes the tunica intima. The thick intermediate layer is the tunica media, the composition of which varies according to the type of vessel and its location. The tunica media of the great arteries, the aorta and the pulmonary artery, are composed of alternating layers of smooth muscle and elastic fibres. The outer supporting tissue layer is the tunica adventitia.

1.2.3 Cardiac embryology

Recent advances in our understanding of cardiac development have shown that while many heart defects appear very different clinically, many have a common embryological pathology. An insight into the developmental basis of these malformations therefore has a strong bearing on the investigation of congenital heart disease. Transcriptional regulation is highly conserved across species (Chen and Fishman, 2000). As a consequence a significant amount of work has been carried out in model organisms, such as mouse and chick. In all vertebrates cardiac development follows a similar general pattern: specification of the cardiac lineage, fusion of the myocardium and endocardium at the embryonic midline to form a simple tubular heart, looping of the tube to the right, chamber specification and development, remodelling of the outflow tract and, finally, the development of the heart valves.

1.2.3.1 Early cardiac specification

The tissues that give rise to the heart are among the first cell lineages to be established, becoming evident when the embryo is undergoing the process known as gastrulation. In humans, this occurs during the third week of development. Gastrulation, gives rise to the three germ layers of the embryo, the endoderm, the mesoderm and the ectoderm. The embryonic plate in humans, initially possessing two layers, is ovoid. In the midline of the long axis of the oval disk is the primitive streak. During gastrulation cells migrate from the upper embryonic layer through the primitive streak giving rise to the three germ layers. The cells that will become the heart, the cardiac progenitors, are among the first mesodermal cells to gastrulate

through the primitive streak (Garcia-Martinez and Schoenwolf, 1993; Schoenwolf and Garcia-Martinez, 1995).

Where the cardiac progenitor cells migrated to from the primitive streak was a matter of some controversy. However recent evidence suggests that the anterior lateral plate mesoderm is the most likely location. In explantation studies, removal of the far anterior lateral plate mesoderm resulted in complete loss of the cardiac structures on the operated side (Ehrman and Yutzey, 1999). Loss of the medial endoderm and mesoderm had no such effect on cardiogenesis (Ehrman and Yutzey, 1999). Once in the anterior lateral mesoderm, the cardiac precursor cells begin to form the primary heart field adopting a horse-shoe shape, termed the cardiac crescent (fig 1.3)

The tissue interactions and signalling molecules that control the earliest stages of cardiogenesis have yet to be fully characterised. Cardiac progenitors become specified in the primitive streak. Signals for this specification are thought to be present within the primitive streak. Non-cardiogenic cells transplanted into the primitive streak are recruited to a cardiac lineage (Schoenwolf and Garcia-Martinez, 1995). The regulatory mechanisms that control the initial recruitment of multipotent cell populations to a cardiac lineage within the primitive streak have proven difficult to elucidate, and are as yet unknown. Once the cardiac progenitors have migrated to the anterior lateral plate mesoderm, there is extensive evidence for a role of the underlying anterior endoderm in cardiac determination and differentiation (Lough and Sugi, 2000). Posterior noncardiogenic mesoderm was recruited to a cardiac lineage following co-culture with anterior endoderm (Schultheiss et al., 1995). Several fibroblast growth factors (Fgf) were found to be expressed in the anterior lateral plate endoderm, and Fgf signalling was demonstrated to be required for cardiomyocte differentiation (Zhu et al., 1996). Inhibition of bone morphogenetic protein 2 (Bmp2), also expressed in the far lateral plate endoderm, blocks differentiation of cardiogenic mesoderm (Schultheiss et al., 1997). In culture, transient exposure of posterior lateral mesoderm to Fgfs and Bmps was sufficient to induce cardiogenesis (Barron et al., 2000). Local application of Bmp-2 was found to give rise to cardiac gene expression in medial mesoderm demonstrating the potential of the anterior medial mesoderm to enter the cardiac lineage given the right environment (Schlange et al., 2000).

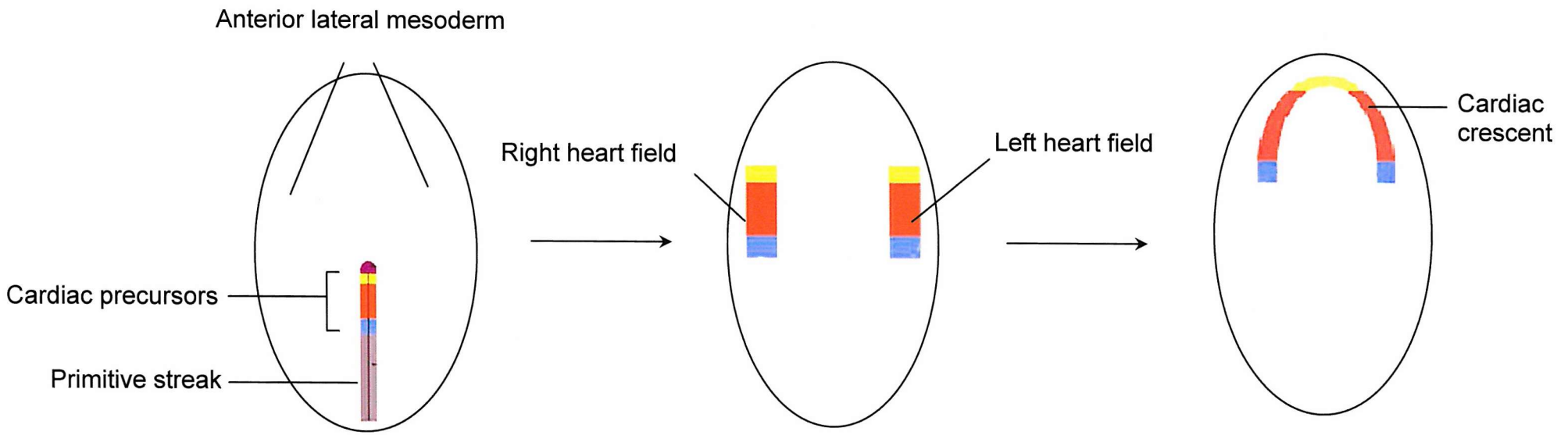

Fig 1.3: Development of the cardiac crescent within the early embryo. The beige ovoid disks represent the embryonic plate. The cardiac progenitors are amongst the first cells to migrate through the primitive streak to the anterior lateral mesoderm where they give rise to the primary bilateral heart fields, and subsequently the cardiac crescent (adapted from Brand, 2003).

Inhibitory factors also play an important role in cardiogenic specification. Regulatory molecules, such as members of the Wnt family, which inhibit cardiogenesis have recently been identified. Ectopic anterior expression of the Wnt8c protein was shown to inhibit cardiac differentiation (Marvin et al., 2001). During normal cardiac development the Wnt inhibitor, Crescent, is expressed in anterior cardiogenic regions of the embryo (Marvin et al., 2001). It was found that the Wntc protein was expressed in the posterior lateral mesoderm, but not in the anterior heart forming region (Marvin et al., 2001), may be due to the inhibitory affects of Crescent.

Cardiac specification therefore seems to be the result of stimulatory and inhibitory factors creating a permissive environment within the anterior lateral plate mesoderm. These inductive signals stimulate the activation of the cardiac transcription programme. Studies in model organisms have demonstrated a role for Bmps in the induction of Nkx2.5, a transcription factor that regulates the expression of many genes that encode cardiac contractile proteins (Schultheiss et al., 1997). Nkx2.5 is one of the earliest markers of avian mesoderm that is fated to give rise to cardiac muscle (Schultheiss et al., 1995). The transcriptional regulators, Gata4 and Tbx5, are also expressed in the anterior lateral plate mesoderm and have been implicated in cardiac lineage determination and differentiation (Srivastava and Olson, 2000). Mutations in all three of these transcription factors are known to be a cause of congenital heart disease (discussed in more detail in section 1.3.3)

1.2.3.2 Formation of the primitive heart tube.

After cardiac mesoderm is specified, the cardiac precursors of the bilateral primary heart fields migrate ventrally, fusing at the embryonic midline to form a single heart tube (see fig 1.4). As the whole embryo undergoes ventral folding, the anterior plate mesoderm is carried with it to the ventral midline. This migration is thought to be controlled to a certain extent by the graded distribution of fibronectin along the cranio-caudal axis. Blocking the interaction between fibronectin and integrin in avian embryos leads to complete or partial failure of the cardiac primordia to fuse, a condition termed cardia bifida (George et al., 1997).

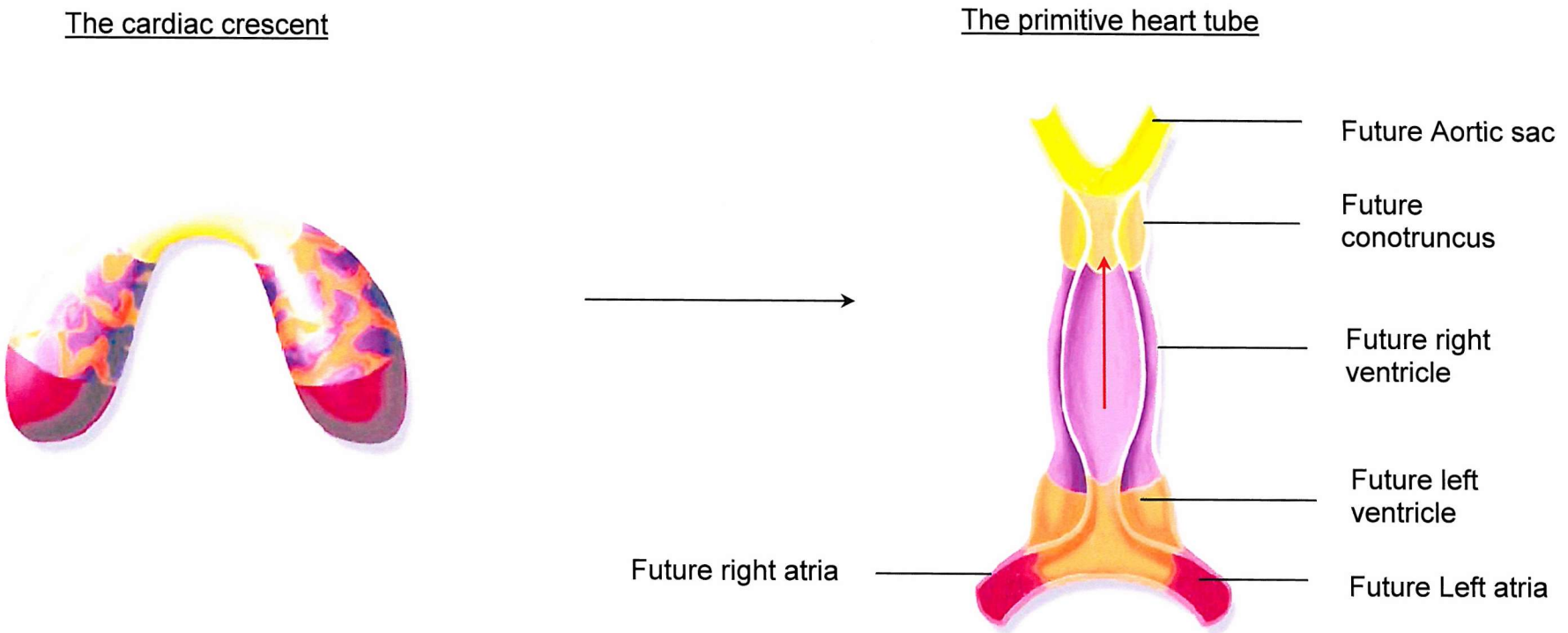

Fig 1.4: Formation of the primitive heart tube. Formation of the primitive heart tube takes place at approximately day 21 in humans. Ventral folding of the embryo brings the arms of the crescent to the midline where they fuse to form the primitive heart tube. At this stage the inflow tract is located caudally. The direction of blood flow is indicated by the red arrow (adpted from Srivastava and Olson, 2000).

Several genes have been identified which are crucial for formation of the single heart tube. Mutations in *Gata4* in mouse gave rise to cardia bifida (Molkentin et al., 1997), as did knock-out studies of the *Gata5* and *Hand2* genes in zebrafish (Reiter et al., 1999). The endoderm is believed to play an important role in cardiac mesoderm migration. Zebrafish that lacked endodermal tissue also developed cardia bifida (Alexander et al., 1999).

The primitive heart tube is comprised of two layers, an inner endocardial lining and an outer myocardial layer. The myocardial and endothelial layers of the heart tube are separated by an acellular extracellular matrix, the cardiac jelly. In humans myocardial contractions are initiated around day 21. The heart tube is centrally positioned within the embryo, located within the pericardial cavity and is bilaterally symmetrical. It remains linked to the overlying dorsal mesoderm through a structure known as the dorsal mesocardium. The two arms of the inflow tract, positioned caudally, are continuous with the developing venous tributaries of the embryo, yolk sac and placenta. The regions of the primitive heart that will develop into the different cardiac chambers are initially arranged sequentially along the craniocaudal axis, with the ventricle located cranially and the atria caudally. Marking studies have shown that the caudal arms of the initial primary tube are fated to become the precursors of the atrial chambers with the cranial stem destined to be the definitive left ventricle (de la Cruz et al., 1977). Chambers or segments of the tube become recognizable by constrictions that demarcate the sinus venosus (the inflow tract), the common atrial chamber, the atrioventricular canal, the ventricular chamber and the outflow tract. Fate mapping studies in the primitive streak have demonstrated that the craniocaudal position at the time of gastrulation is reflected in the craniocaudal position in the primitive heart tube (Garcia-Martinez and Schoenwolf, 1993).

1.2.3.3 Cardiac looping

The linear heart tube bends to the right in a process called looping, the function of which is to displace the atrial myocardium cranial to the ventricular chambers (see fig 1.5). Looping takes place at approximately human embryonic day 23, and in mouse at day 8. The direction of cardiac looping reflects a more global establishment of left/right asymmetry that also affects the other major organ systems of the body.

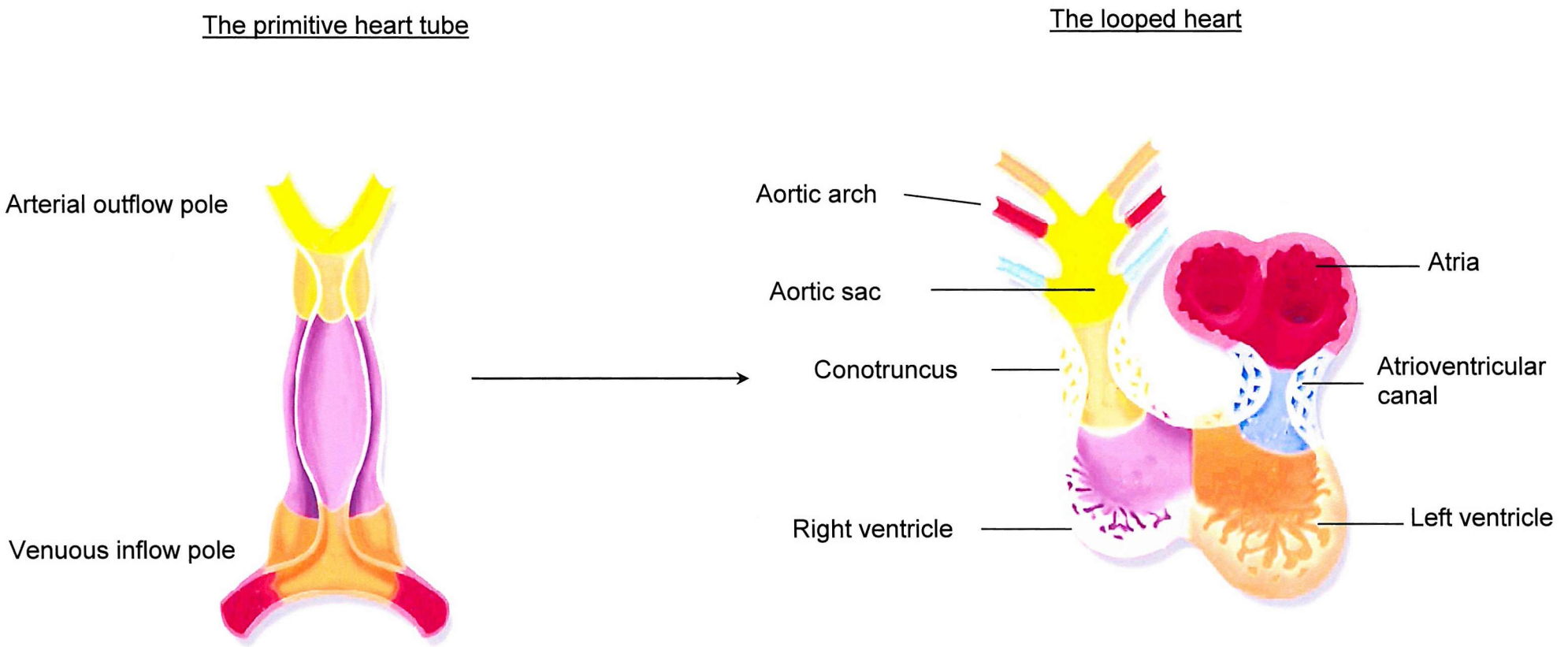

Fig 1.5: Rightward looping of the primitive heart tube. The purpose of looping is to shift the venous inflow pole cranial to the ventricular chambers. Looping of the heart takes place at approximately embryonic day 23 in humans (adapted from Srivastava and Olson, 2000).

A specialised structure, Hensons node, located within the primitive streak is thought to control the establishment of asymentrical LR gene expression throughout the lateral plate mesoderm, which contributes to most of the visceral organs (Levin et al., 1995). Studies have found that, in mice, Hensons node contains ciliary processes which beat, leading to the left-right gradient of morphogens across the embryo (Nonaka et al., 1998). Mutations in the left-right dynein *(Ird)* gene result in immotile cilia, giving rise to *situs inversus* in mice (Supp et al., 1997).

Several regulatory proteins have been shown to play key roles in looping. Nodal, a member of the Tgfβ family, is regarded as the principal signal transducer to the left-sided progenitors of visceral organs (Capdevila et al., 2000). A key target of Nodal signalling in the lateral plate mesoderm and heart is the homeobox gene, paired-like homeodomain transcription factor (Pitx2). The function of Pitx2 is currently unknown, although over-expression in the right lateral plate mesoderm in chick induces the formation of symmetrical unlooped hearts indicating a possible key role in cardiac looping (Logan et al., 1998).

During cardiac looping the primitive heart tube grow dramatically in length through the addition of myocardium to both the venous and arterial poles. Whilst myocardium is added to the venous pole from the primary heart fields, recent studies have identified a novel population of cardiac precursor cells that give rise to the myocardium at the arterial pole. Early experiments raised the possibility that the outflow tract grows by addition of cells derived from mesodermal precursors that lie anterior to the early heart tube (de la Cruz et al., 1977; Viragh and Challice, 1973). More recent studies have confirmed these earlier investigations through identification of a source of outflow tract myocardium from a secondary heart field, distinct to the cardiac crescent, located in the mesoderm adjacent and anterior to the arterial pole of the heart. This has been termed the anterior heart-forming field (AHF). Fate-mapping experiments have confirmed the addition of myocardial cells derived from this precursor population at the arterial pole of the heart (Mjaatvedt et al., 2001). Explant experiments in chick embryos provided further evidence (Mjaatvedt et al., 2001). Removal of the primary heart field resulted in elimination of all cardiac structures with only a beating outlet-like structure being observed. Waldo *et al.* reached similar conclusions, identifying a source of outflow tract myocardium in a secondary heart

field in the anterior mesoderm adjacent arterial pole (Waldo et al., 2001). Members of the Bmp and Fgf protein families have been shown to have a role in AHF recruitment possibly through activation of a similar signalling pathway, involving the transcription factors Nkx2.5 and Gata4, to that operates in the paired heart-forming fields (Waldo et al., 2001). Nkx2.5 is known to directly activate the *Hdf* gene, versican, an important downstream target that is known to be required for normal outflow tract formation (Mjaatvedt et al., 1998). Experiments investigating the role of Fgf10 have also lead to the suggestion that the AHF could give rise to the myocardium of the right ventricle as well as the outflow tract (Kelly et al., 2001; Kelly and Buckingham, 2002). The fact that the interventricular region is a major lineage boundary between cardiomyocytes derived from distinct precursor cell populations may support this.

1.2.3.4 Chamber development

By the end of looping, the heart tube is composed of a single atria and ventricle separated by the atrioventricular canal. Chamber development takes place at the end of the fourth week in humans. Recent studies have shown that the chambers are thought to balloon out from the outer curvature of the looped heart with the development of working myocardium (Christoffels et al., 2000). The primary myocardium of the linear heart is characterised by a low density of gap-junctions and by slow and uniformly spreading conduction velocities of action potentials. This is in marked contrast with the working chamber myocardium where much higher conduction velocities and a higher density of gap-junctions are seen (de Jong et al., 1992; Moorman et al., 1998). The inner and outer curvatures again differ by the presence of trabeculations which are unique to the ventricular outer curvature. The inner curvature remains smooth–walled. These observations led Christoffels *et al.* to examine the expression pattern of several chamber specific genes in mouse and rat embryos (Christoffels et al., 2000). Their findings led to the conclusion that a distinct transcriptional programme operates within the ventral side of the linear heart tube which gives rise to the chamber myocardium of the outer curvature of the looped heart. The chamber myocardium specifically initiates the expression of gap junction genes *connexin 40* (*Cx40)* and *Cx43* (Delorme et al., 1997), as well as *Anf, Chisel* and *Hand1* (Christoffels et al., 2000). Christoffels *et al.* therefore proposed that the ventricular chambers balloon out from the outer curvature in a segmental fashion developing secondary myocardium. The outer curvature becomes the site of active

growth as the chamber myocardium expands, trabeculates and develops a high–velocity conduction system (see fig 1.6). The inflow tract, the AV canal, and the outflow tract myocardium maintain the poor electrical coupling and a long contraction duration that is characteristic of the linear heart tube (de Jong et al., 1992). Several knock out studies have identified genes which are critical for chamber specific development. Ablation of *Mef2c* and *dHand/Hand1* gave rise to a formed primitive heart but failed to form a right ventricular segment suggesting that these factors are required for proliferation and differentiation of the secondary myocardium of the right ventricle (Lin et al., 1997; Srivastava et al., 1997). In *Nkx2.5* knock-out mice all molecular markers of the outer myocardium, Anf, Hand1 Chisel, are abolished or severely downregulated, resulting in the development of an unlooped primitive heart (Lyons et al., 1995; Tanaka et al., 1999).

Defects in chamber development can be a significant cause of peadiatric mortality. Hypoplastic left heart syndrome (HLHS), resulting from under-development, hypoplasia, of the left ventricular myocardium, contributes to 25% of all deaths within the first week of life (Bradley, 1999). It is characterised by a severly reduced left ventricle, and, usually, a diminuitive aorta. It can also be associated with outflow tract obstructions (see section 1.2.3.7). As a consequence, in contrast to the normal fetal circulation, blood is shunted from the left to the right atrium and enters the pulmonary circulation via the right ventricle. Oxygenated blood is re-distributed to the systemic circulation via the ductus arteriosus. On birth, the ductus arteriosus disintergrates, which results in insufficient oxygenated blood entering the systemic circulation. Almost all of affected children will die within 6 weeks of life unless treated. Improvements in surgical repair have had a significant impact on survival, but the prognosis is still poor.

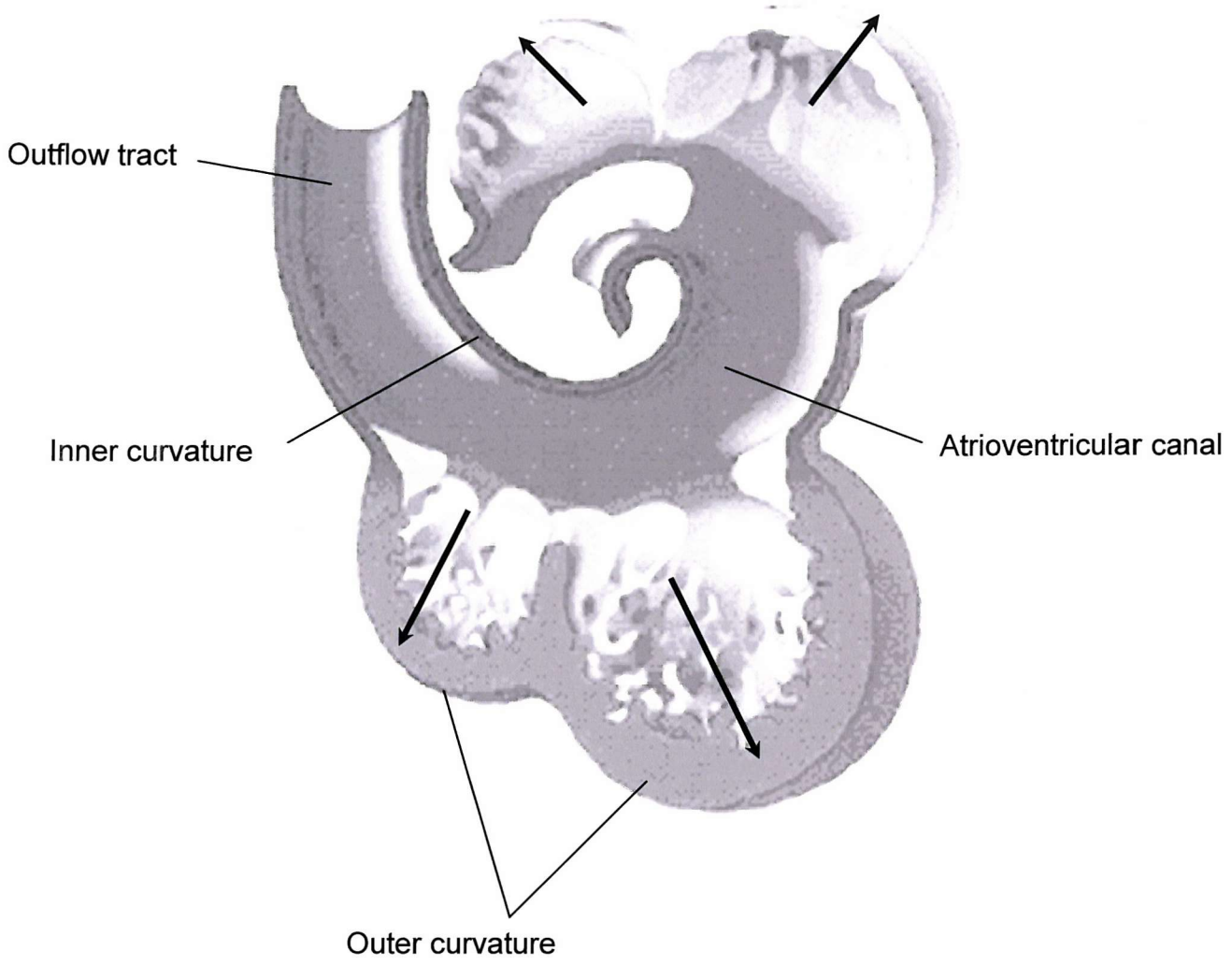

Fig 1.6: Development of the secondary, chamber myocardium. A model has been put forward suggesting that the myocardium of the atria and ventricles balloons out from the outer curvature of the looped heart. Expansion of the ventricular is indicated by the black arrows. The inner curvature retains the characteristics of the primary myocardium of the primitive heart tube (adapted from Christoffels et al., 2000).

1.2.3.5 Septation and remodelling of the atrioventricular canal.

The common atrioventricular canal has to be divided into separate left and right channels. This is achieved through the development of two regional outgrowths, the inferior and superior endocardial cushions (Eisenberg and Markwald, 1995). The cushions are first evident as localised swellings in the cardiac jelly. Subsequent to this expansion, endothelial cells of the AV endocardium are transformed into a mesenchymal phenotype; they then invade the underlying cardiac jelly. Following mesenchymal transformation and invasion of the underlying cardiac jelly, the cushions expand across the common AV canal and fuse to split it into separate left and right channels. In contrast, the endothelium of the ventricles and atria retain their epithelial phenotype. AV septation is completed late in the seventh week.

The mechanisms that regulate cushion placement are unclear. The regionally restricted formation of mesenchyme within the heart tube is thought to be due to the localisation of both myocardial-inducing activity and an endothelial cell competence. Using an *in vitro* collagen gel system culturing embryonic chick heart explants, Runyan and Markwald demonstrated that an inductive signal from the AV myocardium is required to activate endothelial cells to undergo an epithelial-to-mesenchymal transformation (EMT; Markwald et al., 1981). Myocardial induction of the underlying endocardium is thought to be achieved via a secreted factor, the adherons (Markwald et al., 1981; Markwald et al., 1990). These molecular complexes accumulate within the cardiac jelly of the AV canal immediately prior to cushion transformation and have been shown to contain proteins that include fibronectin and transferrin. Following EMT, mesenchymal cells migrate into the underlying cardiac jelly, a process which involves numerous changes in cellular interactions. NCam is a cell adhesion molecule, expressed throughout the endocardium prior to cushion formation. NCam expression decreases dramatically as cells within the cushion-forming regions detach from the endothelial layer before their migration into the cardiac jelly (Crossin and Hoffman, 1991). As NCam is down-regulated, endothelial cells begin to express the, cytotactin, also known as tenascin, which may disrupt cell-substrate adhesions allowing cells to migrate more freely through the extracellular matrix (Crossin and Hoffman, 1991).

Components of the cardiac jelly are suspected to play fundamental roles in cushion development. A major constituent of the cardiac jelly is the glycosaminoglycan hyaluronan (HA). Digestion of HA in rat embryos resulted in a significant reduction in cushion development (Baldwin et al., 1994). In addition embryos lacking the HA-binding proteoglycan versican also fail to develop normal endocardiac cushions (Mjaatvedt et al., 1998). Several genes, such as *Bmp4* and *msx2*, have been identified with expression patterns restricted to the AV canal and outflow tract myocardium, another site of endocardial cushion development (discussed in section 1.2.2.7; Chan-Thomas et al., 1993). Members of the Tgfß family have also been implicated in the formation of cushion mesenchyme. Tgfß1 is initially expressed throughout the cardiogenic mesoderm at high levels in a ubiquitous pattern. By the onset of cell migration into the cushions its expression becomes restricted to the endocardium of the AV canal (Dickson et al., 1993). Treatment of AV explants with Tgfß3 antibodies will block the formation of mesenchyme (Runyan et al., 1992). Tgfß expression is thought to be regulated by the adheron proteins, which has lead to the suggestion that Tgfß serves to elaborate the myocardial signal in endothelially competent cells (Nakajima et al., 1994).

The AV canal also undergoes remodelling to ensure correct chamber alignment. Before septation the ventricles are still in a series formation, with the developing atria communicating only indirectly with putative right ventricle via the left ventricle. Therefore the establishment of a direct communication between right atrium and right ventricle requires not only physical separation of the right and left sides of the initially common atrioventricular junction but also remodelling of the canal myocardium to produce the right AV connection. Recent studies have shown that this is achieved through rightward expansion of the canal ensuring a parallel chamber arrangement (Kim et al., 2001). Subsequent to this expansion, the canal myocardium is incorporated into the smooth walls of the atria just above the AV valves.

Atrioventricular septal defects, also called common atrioventricular canal, can arise from abnormal endocardial cushion development (AVSDs; omim 600309). The defect is a characteristic feature of the Trisomy 21, Down Syndrome (OMIM: 190685) and can be associated with other syndromes, such as 3p25 deletion syndrome (Green et

al., 2000). Isolated cases have been documented as autosomal dominat traits with variable expression and incomplete penetrance (Digilio et al. 1993).

1.2.3.6 Chamber septation

Prior to separation of the primitive atria into a left and right side, the systemic venous vasculature undergoes extensive remodelling to ensure that it enters exclusively on to the right side of the single chamber. Concomitant with this systemic re-orientation, the pulmonary vein develops as a new structure, gaining access to the presumptive left atria through the dorsal mesocardium (Webb et al., 2001). Subsequent to these changes, the primary atrium is separated into two separate chambers through the development of a muscular partition from its dorsal roof, the primary atrial septum (alternatively known as the septum primum; see fig 1.7). In humans this takes place in the fifth and sixth weeks. This muscular shelf grows into the atrial cavity between the systemic and venous openings. The opening between the lower rim of the septum primum and the endocardial cushions is the primary foramen, and allows oxygenated blood to pass from the right to the left atrium. At its leading edge there is a mesenchymal cap. A further mass of extracardiac mesenchyme, the vestibular spine, grows into the heart inferiorly through the rightward border of the pulmonary venous orifice (Webb et al., 2001). There is some debate whether the mesenchymal cap on the leading edge of the primary atrial septum originates from the vestibular spine (Kim et al., 2001) or whether it arises from local epithelial to mesenchymal transformation similar to that seen in the AV cushions (Arrechedera et al., 1987). In the second half of the sixth week with continuing growth of the primary atrial septum, the mesenchymal tissues of the primary atrial septum and the vestibular spine fuse with the endocardial cushions of the AV canal (Kim et al., 2001). This closes the primary foramen. Soon after closure the vestibular spine undergoes myocardialisation. Before fusion with the endocardial cushions takes place, apoptosis creates perforations in the upper portion of the septum that coalesce to form a second interatrial opening, the secondary foramen. The thin upper margin of the primary septum persists as the flap valve of the foramen ovalve. Subsequent to these events a secondary atrial septum develops. This structure forms from infolding of the dorsal atrial wall to the right of the primary septum. On birth, increased left atrial pressure forces the secondary septum against the primary septum closing the foramen ovalve.

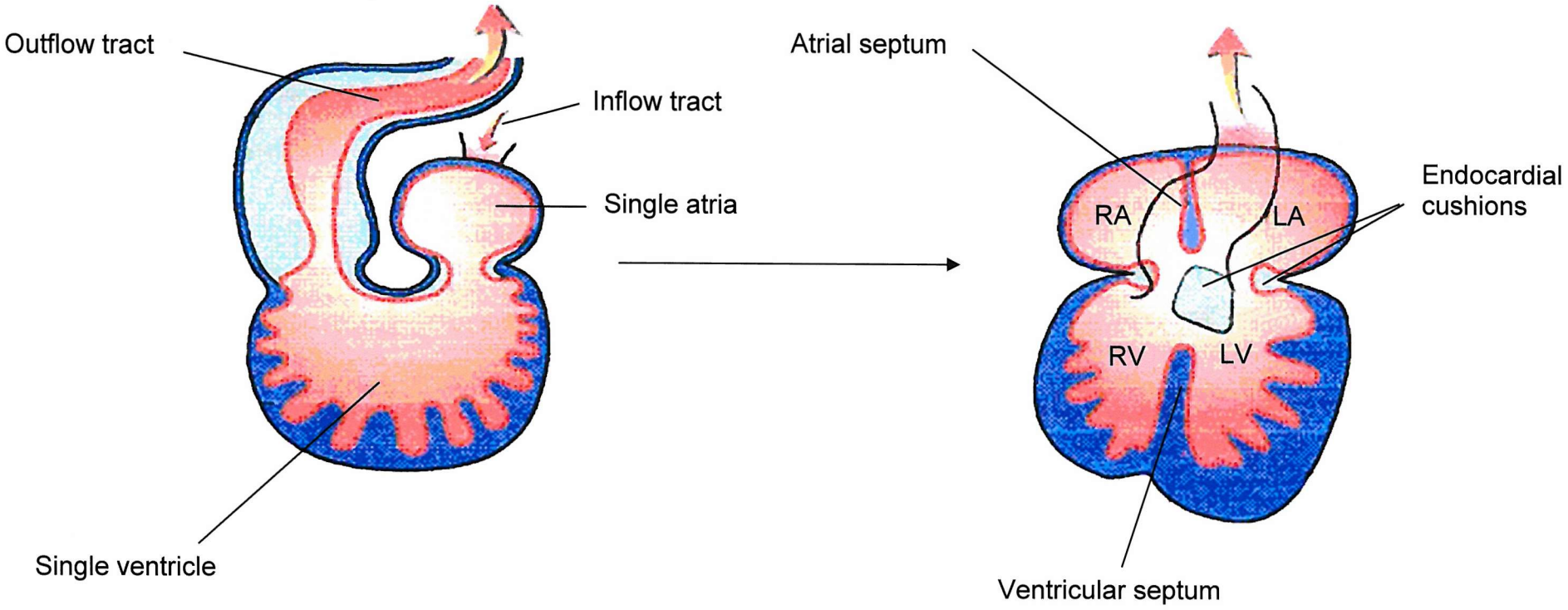

Fig 1.7: Development of the atrial and ventricular septums. The single channel arrangement of the tubular heart is divided into a parallel chamber formation through the development of the atrial septum, the ventricular septum and the atrioventricular canal septum (adapted from Chen and Fishman, 2000). RA: right atria; RV: right ventricle; LA: left atria; LV: left ventricle.

At the same time as the formation of the interatrial septa, the common ventricle is divided into the left and right ventricles, through the development of the interventricular septum (see fig 1.7). The muscular part of the septum is produced concomitant with the ballooning of the left and right ventricles. As the chambers expand out, the septum is formed between them from the deepest convexivity of the ventricular loop and grows towards the atriorventricular septum. For some time there was disagreement concerning the developmental structures that contributed to the ventricular septum. These discrepancies arose due to the differences in development between chick and mammals. In chick it was shown that a muscular AV canal septum contributed to the ventricular septum (De la Cruz et al., 1983). No such structure is apparent in humans with the entirety of the ventricles being separated by the primary ventricular septum, save for a small opening just prior to joining with the AV cushions, the interventricular foramen. Closing of this gap and complete separation of the two ventricles is achieved through fusion of the outflow tract septum with the right side of the interventricular septum (discussed in section 1.2.3.7).

Defects in chamber septation are a common cause of human CHD. Incomplete septation of the ventricles can result in a ventricular septal defect (VSD), the most frequent form of CHD (Hoffman and Kaplan, 2002). Atrial septal defects (ASDs) are not infrequent with the majority resulting from abnormalities in the growth of the secondary atrial septum. Some of these defects may close spontaneously within the first year of life. The size of the septal defect will determine severity. These can be asymptomatic and not diagnosed until adulthood. Many infants have a patent foramen ovalve which can result in a small left-right shunting of blood (Hoffman and Kaplan, 2002). Larger defects may be a significant cause of morbidity.

1.2.3.7 Outflow tract development and septation.

The outflow tract connects the embryonic right ventricle with the aortic sac and can be recognised by the end of the fourth week. The tract is comprised of two components a distal 'truncus' region and more proximal 'conus' portion. The two regions are separated by a characteristic dog-leg bend, which initially serves as a one-way valve to prevent regurgitation of blood back into the ventricle in the absence of the semi-lunar valves. Initially the outflow tract consists almost entirely a of muscular wall separated from the endothelial lining by a sheath of endocardial jelly. Located

within the aortic sac are the pharyngeal or aortic arches. Two of these, the fourth and sixth, will connect the separating outflow tract with the dorsal aorta and pulmonary artery respectively.

The purpose of septation is to separate the single tract into two vessels, the aorta and the pulmonary artery. This begins in the course of the fifth week with the development of two opposing enodcardial cushions which expand to create the septal and the parietal ridges. The formation of the ridges begins proximally and moves into the distal outflow tract following a rightward spiralling course to the aortic sac. Within the aortic sac itself a transverse wedge of mesenchymal tissue develops to separate the arteries feeding the fourth and sixth aortic arch. This wedge is termed the aorto-pulmonary septum. Fusion of the endocardial ridges takes place as the ridges expand across the outflow tract lumen, proceeding into a distal to proximal direction. The distal ends of the joined cushions fuse with the wedge-shaped mesenchyme of the aorto-pulmonary septum. This results in connecting the developing aorta with the artery of the left fourth arch, which becomes the aortic arch. At the same time the pulmonary trunk is linked with the left sixth arch, which will upon birth link through to the lungs. Concomitant with septation of the distal outflow tract significant changes take place with the walls of the arteries. Rapid regression of the myocardium from the distal portion towards the bend takes place as the walls take on an arterial phenotype (Ya et al., 1998). The abrupt shortening of the myocardial portion of the outflow tract has been previously described as the 'absorption of the conus' (Watanabe et al., 1998). The cardiomyocytes show no signs of apoptosis which suggested that the cardiomyocytes may undergo transdifferentiaition (Ya et al., 1998).

By 7.5 weeks the endocardial ridges of the proximal outflow tract have fused. However in addition to the two endocardial ridges, two further intercalated cushions have grown in the opposite quadrants of the outflow tract. These give rise to the primordia of the arterial valves and are found in the most distal section of the proximal region (development of the valves is discussed in the next section). In marked contrast to the distal outflow tract, which saw a regression of myocardial tissue, muscular tissue is now added to the outflow tract in its most proximal portion, in a process known as 'myocardialistion' (Ya et al., 1998). The myocardial cells in the walls grow into the most proximal parts of the cushions as they fuse, converting the

endocardial septum into a muscular partition. As the cushions fuse and muscularise, they span the outlet from the right ventricle, which initially supports both outflow tracts. At this stage the aorta is in continuity with the developing right ventricle and can only communicate with the left ventricle through the interventricular foramen. To unite the aorta exclusively with the left ventricle, the leading edge of the fused cushions expands across the roof of the right ventricle and attaches to the right ventricular surface of the crest of the interventricular septum, effectively walling the aorta into the left ventricle and closing the interventricular foramen.

Cells derived from the neural crest have been known for some time to have a role in outflow tract septation (Le Lievre and Le Douarin, 1975). Their exact role has been extensively studied and is still a matter of some debate. The cardiac neural crest invade the distal endocardial ridges and can be traced as rods of densely staining mesenchymal tissue within the cushions penetrating as far as the dog-leg bend and beyond into the proximal region. Although direct evidence is lacking the neural crest are suspected to have a crucial role in fusion of the ridges. Cardiac neural crest ablation in chick embryos resulted in persistent truncus arteriosus (PTA), a condition characterised by an undivided outflow tract (Kirby et al., 1983). It is thought that the invasion of neural crest may trigger the onset of fusion, as these cells were shown to be prominent at fusion, but subsequently disappeared (Jiang et al., 2000; Poelmann et al., 1998).

Defects in outflow tract septation can result in persistent truncus arteriosus (PTA) in humans, giving rise to an undivided outflow tract. Aberrant outflow tract (OFT) septation may also result in VSDs and a condition called double outlet right ventricle (DORV) with both outlet channels connected solely with the right ventricle. This may arise due to the failure of the OFT to grow across the roof of the right ventricle and to wall the aorta into the left ventricle and close the interventricular foramen. Defects in vessel alignment can also result in transposition of the great arteries (TGA) with the aorta arising from the right ventricle and the pulmonary artery from the left. Obstructive lesions, stenosis (narrowing), and complete blockage (atresia), of the outflow tract vessels can also be a cause of CHD. Isolated pulmonary stenosis represents 7.8% of all cardiovascular malformations, and is associated with Noonan syndrome (OMIM: 163950; Marino and Digilio, 2000). Left-sided obstructions can

include supravalvular aortic stenosis (SVAS) and aortic co-arctation. SVAS consists of a narrowing of the aorta immediately above the valve, and is a consistent feature of Williams' Syndrome (OMIM: 1855001.3.3; discussed in more detail in section 1.4.3). Aortic co-arctation is found distally and presents 4.6% of all cardiac malformations (Marino and Digilio, 2000).

1.2.3.8 Development of the heart valves

The last stage in heart development is formation of the cardiac valves. The valves develop from the endocardial cushions within the AV canal and outflow tract. Correct development of the endocardial cushions, as discussed in section 1.2.2.5, is therefore central to valvulogenesis. Initially mesenchymal tissue derived from the fused AV cushions serves as rudimentary valves before definitive mitral and tricuspid valves are formed. The cushion tissue is thought to transdifferentiate into the fibrous connective tissue of the valve leaflets (Lamers et al., 1995). The extracellular matrix protein fibulin-1 is thought to play a role in the differentiation of the cushion mesenchyme into the fibrous structure of the valve (Spence et al., 1992) through proving a scaffold for cell adhesion. Other ECM molecules have been implicated as having a role in valvulogenesis (Schroeder et al., 2003). Knock out studies have identified some of the key transcriptional regulators of valve development. Nfatc, is expressed specifically in the forming embyronic valves. Deletion of *Nfatc* in mice results in absence of cardiac valve formation (Ranger et al., 1998). Ablation of *Smad6*, another crucial valvular specific transcription factor, gave rise to hyperplastic valve tissue (Galvin et al., 2000).

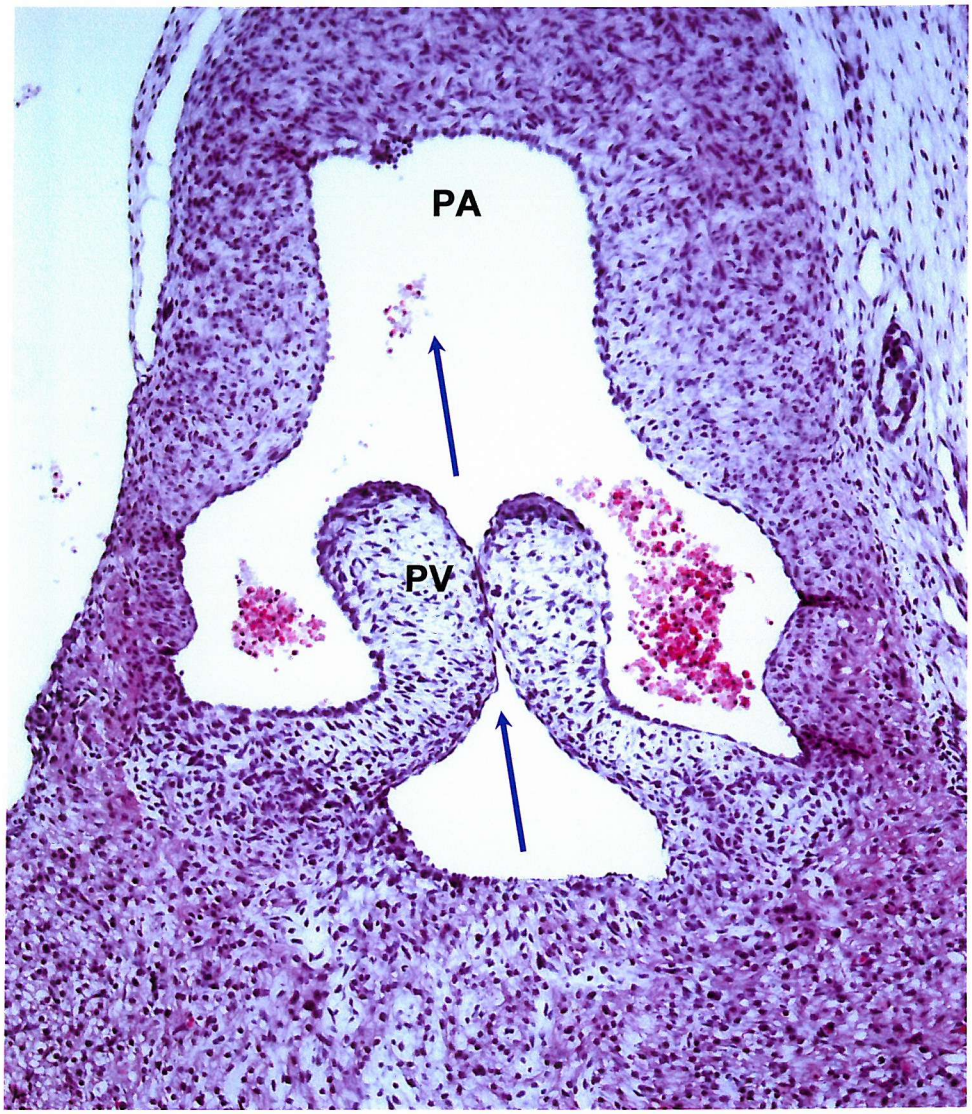

Fig 1.8: A heamatoxylin and eosin stained section showing the pulmonary valve of an 8 week human fetal heart. The valves are the last structures to develop. The semi-lunar valves, the pulmonary and aortic valves are located within the outflow tract. The atrioventricular valves, the bicuspid and the mitral valves are located between the right-sided and left-sided chambers respectively. During right ventricular contraction de-oxygenated blood (indicated by the blue arrows) is forced through the pulmonary valve into the pulmonary artery and on to the lungs to re-oxygenate. Following contraction, the valve leaflets close, preventing the back flow of blood into the right ventricle

Cardiac development event	Defect that can arise at each developmental stage
• Formation of the cardiac crescent within the anterior lateral mesoderm	
• Migration of the cardiac precursors to the ventral midline and fusion to form the primitive heart tube	• Cardia bifida (experimental)
• Rightward looping of the heart.	
• Development and ballooning of the chamber myocardium	• Hypoplasia and hyperplasia of the cardiac chambers
• Development of the endocardial cushions and septation of the atrioventricular canal. Rightward shifting of the AV canal to ensure correct chamber alignment	• AVSDs • Common atrioventricular canal • Double inlet left ventricle
• Growth of the atrial and ventricular septums	• VSDs • ASDs
• Outflow tract alignment and septation	• Truncus arteriosus • Transposition of the great arteries • Double outlet right ventricle • VSD • Coarctation of the aorta • SVAS
• Development of the atrioventricular and semi-lunar valves	• Valvular atresia (blockage) and stenosis (narrowing)

Table 1.1: The congenital heart defects that can arise at each developmental event. AVSD – atrioventricular septal defects; VSD – ventricular septal defect; ASD – atrioventricular septal defects; SVAS – supra valvular aortic stenosis

1.3 A genetic aetiology to congenital heart disease

Until relatively recently, the predominant view was that the causes of congenital heart disease were largely multifactoral involving interactions between genes and environment rather than due to mendelian inheritance of single gene mutations (Nora, 1993). Environmental factors can play a role in the aetiology of congenital heart disease. Known teratogens, such as insulin-dependant maternal diabetes and maternal phenylketonuria, have both been associated with an increased risk of cardiac defects in their offspring (Ferencz et al., 1990; Levy et al., 2001). A recent study found that over-weight women were more likely than those of average-weight to have children with heart defects (Watkins et al., 2003). Studies have also demonstrated that correct functional loading of the heart is necessary for normal cardiogenesis in chick embryos (Sedmera et al., 1999). Disruption to normal blood flow could therefore be a cause of human CHD. The mitral and aortic valve atresia/stenosis seen in hypoplastic left heart syndrome may disrupt blood flow preventing correct development of the left chamber myocardium. Although in the majority of cases no single underlying cause can be identified, it has become increasingly clear that there is evidence for a genetic aetiology to some forms of congenital heart disease. The evidence is threefold:

1. The increased recurrence risk seen in offspring and siblings.
2. The significant association of cardiac defects with chromosomal abnormalities.
3. The identification of pathogenic mutations in cardiac developmental genes in humans.

1.3.1 The recurrence risk

Improved survival rates of patients with major congenital heart defects have enabled increased numbers to live to adulthood and parenthood. Analysis of recurrent congenital heart defects in the offspring of such individuals has provided valuable clues into the aetiology of CHD. The approximate risk of CHD within the population is 0.7% (Gill et al., 2003). Several studies have indicated an offspring and sibling recurrence risk significantly higher than seen in the general population (Gill et al., 2003; Burn et al., 1998). In 1998, Burn *et al.* calculated the offspring recurrence rate to be 4.1%, and the sibling recurrence rate to be 2.1%. The offspring recurrence risk

was significantly higher if the mother was affected, 5.7%. The offspring recurrence risk for a male proband was 2.2%. Variations in risk were also associated with different cardiac malformations (see table 1.2). For example, for parents with an atrioventricular septal defect there was an offspring risk of 7.8%. A recent investigation into a large cohort of patients with bicuspid aortic valve (BAV), calculated that, in this population, determination of BAV was almost entirely genetic (Cripe et al., 2004)

Numerous reports of familial occurrence with multiple affected members have been described (Garg et al., 2003; Schott et al., 1998). Many follow a mendelian inheritance pattern, consistent with an autosomal dominant and autosomal recessive inheritance patterns. Such cases argue strongly for a monogenic aetiology, with disruption to a single crucial cardiac developmental gene.

1.3.2 Chromosomal abnormality

Morphogenesis requires that certain gene products are present in the appropriate 'dose'. On the autosomal chromosomes, one copy of each gene is inherited from each parent, giving a normal dose of two copies. Loss or gain of chromosomal material, resulting in aberrant gene dosage, can have profound implications for embryonic and fetal development. Several studies have shown an association between chromosomal abnormality and congenital heart disease, the most comprehensive to date being the Baltimore Washington Infant Study (BWIS). Their findings showed that 13% of CHD children had an abnormal karyotype (Ferencz et al., 1989). Chromosomal abnormality is also a major cause of pre-natal death. In 1999 a study sought to investigate the frequency of chromosomal aberrations amongst fetuses with CHD sent for necroscopy (Tennstedt et al., 1999). Over a seven year period necroscopy was carried out on 815 fetuses. A congenital heart defect was present in 129 cases. Cytogenetic analysis detected chromosomal abnormalities in 43 cases (33%) of the CHD cohort. However these may be underestimates of the contribution of chromosomal abnormality to congenital heart defects, due to the limited ability of current routine diagnostic tests to identify microdeletions.

Defect in proband	Affected births/live births among offspring		Total risk
	Of fathers	Of mothers	
Tetralogy of fallot	2/124 (1.6%)	6/132 (4.5%)	3.1%
TGA	0/14	0/6	-
Abnormal connections	1/22 (4.5%)	2/17 (5.9%)	5.1%
Atrioventricular septal defects	1/13 (7.7%)	3/38 (7.9%)	7.8%
APVC	0/10	1/17 (5.9%)	3.7%
All heart defects	4/183 (2.2%)	12/210 (5.7%)	4.1%

Table 1.2: The recurrence rate of different cardiac malformations. In all cases the recurrence risk for offspring of affected mothers was higher than that seen when the father was the proband. TGA-transposition of the great arteries; APVC – Anomalous pulmonary venous connection. (Table taken from Burn, Brennan et al. 1998)

Cytogenetic banding has a resolution that, at best, can only identify deletions of greater than 3Mb, and as a result smaller submicroscopic microdeletions may be going undetected. The introduction of microarray-comparative genomic hybridisation, a novel high-resolution, high-through-put deletion detection technology, may have a significant impact on the identification of submicroscopic deletions (Solinas-Toldo et al., 1997; discussed in more detail in chapter 4). Chromosomal abnormalities may therefore be found to contribute more to CHD than is currently thought.

A recent meta-analysis carried out by Brewer *et al.* sought to construct a genome-wide chromosome map of autosomal deletions associated with 47 congenital malformations, using detailed clinical and cytogentic information on 1,753 patients (Brewer et al., 1998). There were strong associations with numerous cytogentic bands and a variety of cardiac defects. Two chromosomal regions in which there was a highly significant association ($p<.001$) with cardiac malfomations were 22q11 and 11q23-25, deletions in which cause DiGeorge Syndrome (DGS: OMIM-188400), and Jacobsen syndrome (JBS, OMIM 147791) respectively. DiGeorge Syndrome is a clinically heterogenous disorder characterised by cardiovascular defects, thymic, parathyroid and craniofacial anomalies. It occurs with a frequency of 1 in 4000 live births. There is considerable phenotypic variability, which can range from life threatening cardiac abnormality to mild craniofacial disorder and developmental delay. Cardiac defects account for the main cause of mortality and morbidity, and occur in 75% of patients with a deletion. Approximately 80% of DGS cases are caused by a either a 1.5 or a 3 Mb deletion within 22q11 (McDermid and Morrow, 2002).

Jacobsen syndrome is caused by terminal deletions of the long arm of chromosome 11. Patients can present with a range of cardiac phenotypes which include ventricular septal defect, aortic coarctation, and hypoplastic left heart. Extracardiac defects such as triogonocephaly, and thrombocytopenia are also possible. It is likely that the cardiac defects are of variable penetrance, with hypoplastic left heart representing the more severe end of the spectrum. The presence of fragile sites with in this region has been proposed as a possible mechanism for deletion (Jones et al., 2000).

1.3.3 Single gene-defects causing congenital heart disease.

Convincing evidence for a genetic role in congenital heart disease has come from the discovery of mutations in genes involved in cardiac development (table 1.3). For example, mutations in *TBX5* and the elastin gene can cause Holt-Oram and Williams' syndrome respectively.

Holt-Oram syndrome (HOS: OMIM-142900) is an autosomal dominant disorder that causes abnormalities of the heart and upper limbs. The most frequently found cardiac abnormalities include secundum atrial septal defects (ASD), ventricular septal defects (VSD) and tetralogy of Fallot (TOF). Tetralogy of Fallot is a combination of cardiac defects, which include VSD, pulmonary stenosis, overriding of the ventricular septum by the aorta and right ventricular hypertrophy. HOS occurs with an estimated frequency of 1 in 100, 000 live births. Several groups mapped the HOS locus to a region of chromosome 12 (Bonnet et al., 1994; Terrett et al., 1994). The region was further refined through the detailed FISH mapping of a HOS patient with a complex chromosomal rearrangement involving chromosome 12 (Li et al., 1997b). The T-box gene, *TBX5*, was found to map close to the translocation breakpoint and subsequent analysis found mutation in affected individuals in three familial and three sporadic families (Li et al., 1997b).

Supra valvular aortic stenosis (SVAS; OMIM-185500), is an inherited obstructive vascular disorder that causes a significant narrowing of the aorta (Eisenberg et al., 1964). SVAS can occur sporadically, inherited as an autosomal dominant trait or as a feature of Williams' syndrome (OMIM: 194050). The incidence of familial and syndromic SVAS is estimated to be 1 in 20,000 live births (Ewart et al., 1993b). SVAS was mapped to chromosome to 7q11.23 (Curran et al., 1993; Ewart et al., 1993b). Curran *et al.* subsequently identified a balanced translocation t(6:7)(p21.1;q11.23) that disrupted the elastin gene in a family with SVAS (Curran et al., 1993). The breakpoint was localised to exon 28. Li *et al.* carried out mutational analysis which revealed elastin point mutations in four familial and three sporadic cases of SVAS (Li et al., 1997a).

Disorder	Cardiac Pathology	Gene	Function	Reference
Syndromic:				
Alagille syndrome	Pulmonary stenosis, TOF	*JAG1*	Ligand for Notch receptor	(Li et al., 1997b)
Char syndrome	Patent ductus arteriosus	*TFAP2B*	Transcription factor	(Satoda et al., 2000)
Holt-Oram syndrome	ASD, VSD and TOF	*TBX5*	Ttranscription factor	(Li et al., 1997c)
Marfan Syndrome	Aortic and valvular defects	*FBN1*	Fibrillin, connective tissue component	(Dietz et al., 1991)
Williams Syndrome	Supravalvular aortic stenosis (SVAS)	*ELN*	Elastin, principal component of the elastic fibre	(Li et al., 1997a)
Non-syndromic:				
	ASD, VSD, DORV, TOF, HLH	*NKX2.5*	Transcription factor	(Benson et al., 1999; Goldmuntz et al., 2001)
	ASD, VSD, AVSD, valve abnormalities	*GATA-4*	Transcription factor	(Garg et al., 2003)

Table 1.3: Genes in which mutations are known to cause congenital heart disease

Ewart *et al.* found that larger deletions within 7q11.23 gave rise to Williams' syndrome, which in addition to the cardiac anomalies gives rise to neurobehavioural defects and mental retardation (Ewart et al., 1993a). Haploinsufficiency of the elastin gene gave rise only to the SVAS phenotype.

Mutations in the cardiac homeobox transcription factor, *NKX2.5,* have been identified in families with inherited autosomal dominant ASD and atrioventricular conduction block (Benson et al., 1999). Other congenital heart anomalies, including ventricular septal defects, tetralogy of fallots (TOF) and aortic stenosis, were also observed in these families. *NKX2.5* mutations were found in 4% of individuals with non-syndromic tetralogy of fallots (Goldmuntz et al., 2001). A recent study identified a mutation in *NKX2.5* in a patient with ASD in association with hypoplastic left heart (Elliott et al., 2003). The change was found in all affected family members. The variation, causing an amino-acid change from threonine to methionine, took place within the *NKX2.5* homeodomain and was predicted to disrupt its ability to bind DNA. This was the first documented case of mutations of *NKX2.5* being linked with HLHS. A further 18 patients with sporadic or familial HLHS who had a normal karyotype were investigated for NKX2.5 mutations but none were found. Mutations in the *GATA-4* gene were recently identified in a large family, all of whom had ASDs (Garg et al., 2003). Eight individuals had additional forms of congenital heart disease, including VSDs, atrioventricular septal defects, pulmonary valve thickening or insufficiency of the cardiac valves (see section 1.4.1).

1.4 A positional candidate strategy for identifying congenital heart disease genes.

A positional candidate approach has proven to be an effective strategy for identifying disease genes. This first involves mapping the disease to a chromosomal region to establish a defined critical disease interval. This can be achieved either through linkage anlysis of affected families or investigation of patients with chromosomal aberration. Genes mapping to the disease interval can be assessed for candidacy. The strongest candidate genes are investigated to demonstrate a role in disease aetiology

1.4.1 Defining a critical disease region.

Several approaches can be used to define an initial critical region. Linkage analysis requires studying large pedigrees with a high number of disease affected individuals to identify allelic association with the disease phenotype. The recent discovery of mutations in the *GATA-4* gene was facilitated through a linkage analysis approach. Garg *et al.* identified a large kindred with multiple affected members that was consistent with an autosomal dominant inheritance pattern (Garg et al., 2003). Linkage analysis mapped the disease to chromosome 8p22-23 defining a critical region of approximately 12.7 Mb. Mutations were detected in the *GATA-4* gene in all 16 affected individuals. The principle limitation of linkage studies is the need for a large family pedigree with multiple affected members. Whilst a higher number of CHD patients are surviving to reproductive age, there is still a paucity of large families available.

Mapping of chromosomal abnormalities has proven to be highly effective in defining critical CHD regions. Balanced translocations and inversions, which result in no DNA loss, can cause loss-of-function if it disrupts the coding sequence of a gene, or separates it from a nearby regulatory element. The breakpoint therefore provides a valuable clue to the exact physical location of the disease gene. As discussed previously, FISH analysis of a chromosomal translocation aided the identification of the *TBX5* gene and the elastin gene as causes of Holt-Oram syndrome and SVAS respectively (see section 1.3.3). Deletion analysis of affected individuals has also proven to be an effective means of defining critical regions. The disease phenotype will be expected to be the result of haploinsufficiency of a gene or genes within or

near the deletion. If the deletion is large, further reduction of the critical region can be attempted through screening similarly affected individuals for patients with overlapping or smaller nested deletions. For example deletion studies have been used to define critical regions in two chromosomal loci, 10p and 11q23.

Monosomy 10p gives rise to a DiGeorge like phenotype. Early deletion studies mapped the disease locus to a subgenomic region of 10p (Schuffenhauer et al., 1998). It was recently shown that loss-of–function mutations of the *GATA3* gene gave rise to the hypoparathyroidism, sensorineural deafness and renal anomalies seen in these patients (Van Esch et al., 2000). To identify the genetic cause of the cardiac defects seen in $10p^-$ syndrome, Lichtner *et al* established a P1–derived artificial chromosome (PAC) contig spanning the DGCR2 locus (Lichtner et al., 2002). Characterising the proximal and deletion breakpoints in several patients defined a critical region of approximately 300kb. *BRUNOL3* was the only gene identified within this genomic region.

A similar strategy was adopted to define an HLH susceptibility locus within 11q24-ter. An initial 20Mb minimal deletion region associated with the HLH phenotype had originally been defined (Penny et al., 1995). To reduce the critical interval Phillips *et al* undertook molecular analysis, using loss-of-heterozygosity analysis and fluorescence *in situ* hybridisation, of four CHD deletion patients. These analyses permitted the critical region to be reduced to approximately 9Mb, the size of the smallest deletions. A further seventeen HLH patients were investigated with a view to identifying an overlapping deletion to further reduce the critical region, but none were found.

1.4.2 Prioritisation of candidate genes

Genes mapping to a critical region are assessed for disease candidacy. In doing so, it is necessary to take into account that a translocation or deletion breakpoint may disrupt *cis*-acting regulatory elements that influence genes some distance from the breakpoint. This is termed position affect, and has been documented in several diseases, for example campomelic dysplasia (Pfeifer et al., 1999). Position effects as large as 950kb from the disease gene, SOX9, have been reported in this disorder

(Pfeifer et al. 1999). This demonstrates that genes within approximately 1 Mb either side of a breakpoint must be considered as possible candidates for a disease. Candidacy is assessed on known expression and known or predicted function. Knowledge of a knock-out phenotype may also provide further evidence for investigating a candidate gene.

1.4.2.1 Expression analysis

It would be predicted that a gene, mutations in which cause CHD, would be expressed in the heart during development. For example Phillips *et al.* (2002) identified the Junction Adhesion molecule, (*JAM3*), as a candidate gene for the hypoplastic left heart phenotype associated with Jacobsen syndrome (Phillips et al., 2002). This was partly based on expression profiling of *JAM3* using mRNA tissue in *situ* hybridisation within human fetal heart sections. Expression was detected initially in the outflow tract and subsequently in the developing atrioventricular valves, the aortic and pulmonary valves and the pulmonary artery. The valvular expression localised to the endothelial cells that lined the lumen, whereas the arterial expression was within the walls of the aorta and pulmonary artery. Some expression was detected within the ventricular myocardium. This was an expression pattern that fitted well with the valvular atresia/stenosis and ventricular hypoplasia seen in HLH patients and confirmed *JAM3* candidacy.

However the expression of a candidate gene does not have to be restricted to affected tissues. As discussed previously Holt-Oram syndrome affects the heart and upper skeletal limbs. *TBX5* was identified as the disease gene. Analysis of TBX5 expression found high levels in the heart and developing fore limbs, but also the trachea, lung and thoracic wall, structures which are not affected by *TBX5* loss-of-function (Li et al., 1997b)

1.4.2.2 Functional analysis

Known or predicted function can provide further evidence for candidacy. Functional studies of the connective tissue protein, fibrillin-1, suggested a possible role in Marfan syndrome (MFS: OMIM 154700). Fibrillin-1 is a principal component of the microfibrils that are ubiquitous in the connective tissue and make up a large part of the elastic fibres within blood vessel walls. The Marfan disease phenotype, lax joints,

lense dislocation and cardiovascular abnormalities, indicated a possible connective tissue abnormality. Marfan syndrome was mapped to chromosome 15q15-21.3 (Dietz et al., 1991b). The fibrillin-1 gene was cloned and mapped to chromosome 15q21.1 by *in situ* hybridisation (Magenis et al., 1991). Fibrillin-1 was an obvious candidate choice. *De novo* mutations were subsequently identified in two MFS patients (Dietz et al., 1991a).

Knowledge of homologous genes can provide evidence for gene candidacy. For example, subsequent to mutations being found in fibrillin-1, a second fibrillin gene, fibrillin-2, was identified and mapped to 5q23-31. Congenital contractural arachnodactyly (OMIM 121150), a phenotypically similar disorder to Marfan Syndrome, was mapped to 5q. *FBN2* gene was shown to be the causative gene (Putnam et al., 1995). Genes which function within the same biochemical or developmental pathways as known disease genes may also be legitimate candidate targets. This has been highlighted by the finding of mutations in the three cardiac transcription factors; *NKX2.5, GATA 4* and *TBX5*. GATA 4 and TBX5 were previously shown to interact with NKX2.5 (Durocher et al., 1997). A more recent study identified a novel interaction between GATA4 and TBX5 (Garg et al., 2003). Given the similarity in septal defects seen in these patients it has been suggested that these three proteins may form a transcriptional complex to regulate genes involved with cardiac septation (Packham and Brook, 2003).

1.4.2.3 Transgenic knock-out technology

Knock-out studies in model organisms to ablate gene or protein function may give an indication of a potential human disease phenotype. Transgenic knock-out technology was instrumental in identifying the Tbox gene, *Tbx1* as a candidate for DiGeorge Syndrome (Lindsay et al., 2001). More than 40 genes map to the DiGeorge deletion interval. Targeted deletion of the mouse region homologous to 22q11 identified *Tbx1* as the only haploinsufficient gene that cause a DGS-like phenotype (Lindsay et al., 2001). Mice heterozygous for a null mutation in *Tbx1, (Tbx1*$^{+/-}$), developed outflow tract defects similar to those seen in DiGeorge patients (Merscher et al., 2001) . Patients with a DGS phenotype but without a deletion 22q11 have been investigated extensively for mutations in TBX1, but known have as yet been found. Gong *et al.*

reported sequence changes in a small number of patients with a DGS phenotype but the functional significance of these are unclear (Gong et al., 2001).

1.4.3 Confirmation of a CHD gene.

Confirmation that a strong candidate gene is the disease gene is obtained by finding the presence of pathogenic mutations in affected individuals. Identified changes would be expected to disrupt gene or protein function and be absent from a large control panel of normal chromosomes. The main classes of DNA mutations include deletions (ranging from 1bp to megabases), duplications and single base pair substitutions. Deletions, duplications and insertions can give rise to frameshift changes potentially resulting in incorrect amino acid coding. Amino acid changes in functionally important domains or those that are highly conserved between species may be pathogenic. Also incorporation of a different type of amino acid, such as one that differs in charge, may have a deleterious affect on the tertiary structure of the protein. Frameshift changes can also give rise to a premature stop codon resulting in an unstable mRNA or a truncated protein. Single base pair substitutions can give rise to missense mutations, replacing one amino acid with another, or nonsense mutations, replacing an amino acid codon with a stop codon. Mutations identified on analysis of TBX5 in Holt-Oram patients all resulted in the introduction of premature stop codons either directly or through a frame-shift changes which brought a stop codon in frame later in the sequence (Li et al., 1997b).

It is also possible that base pair substitutions that give rise to no amino acid change or that lie within intronic sequence may be deleterious if they affect crucial regulatory elements. Mutations in the consensus sequences that govern exon splicing are important cause of genetic disease (Cartegni et al., 2002; Faustino and Cooper, 2003). 15% of point mutations that give rise to human disease are the result of RNA splicing defects (Krawczak et al., 1992). This is now thought to be an underestimate of their total contribution (Cartegni et al., 2002; Faustino and Cooper, 2003). Most splicing mutations directly affect the consensus splicing recognition sequences, typically leading to skipping of the neighbouring exon, such as those identified in SVAS patients (Li et al., 1997a). Less frequently, mutations can create an ectopic splice site. Activation of a cryptic splice site was identified in the calpain 3 gene of a limb girdle

muscular dystrophy patient (Richard and Beckmann, 1995). No amino acid change was predicted. The aberrant splicing resulted in loss of coding sequence and the introduction of a frameshift. Translationally silent amino acid changes must therefore be scrutinised to ensure that they do not interrupt important regulatory elements.

1.5 Aims

Abnormalities within chromosomal region 1q42-44 give rise to developmental anomalies which can include variable cardiac malformations. CHD associated with monosomy of this region has been reported to be fatal. Such association argues strongly for the presence of a crucial cardiac developmental gene within this region, haploinsufficiency of which can cause congenital heart disease. The aim of the project was therefore to identify and investigate a strong candidate gene through a positional candidate strategy.

The project has three components:

1. Definition of a critical cardiac disease region through molecular investigation of patients with congenital heart disease (chapter 3).

2. Development and validation of micro-array comparative genomic hybridisation for detecting deletions within 1q42-44 in CHD patients (chapter 4).

3. Assessment and investigation of genes within the critical region for CHD candidacy. Undertake mutation analysis of the strongest candidate gene within a panel of CHD patients (chapter 5)

Chapter 2: Methods and Materials

The contents of general lab solutions and buffers made in-house are detailed in section 2.8. Source of materials can be found in section 2.9. All primer sequences and antibodies are listed in appendix I.

2.1 Generic protocols

2.1.1 Cell culture

Fibroblast and transformed lymphoblastoid and cell lines were cultured in RPMI containing glutamine and 10% fetal calf serum at 37°C and at 5% CO_2 in a Heto Cellhouse 170 Hi.

2.1.2 DNA extraction from cultured cell pellets.

DNA was extracted from cell lines using Nucleon Bacc1 DNA extraction kits and protocols (Nucleon kit, Tepnel Life Sciences, Manchester, UK). Cultured cells were pelleted by centrifugation at 1,000rpm for 5 minutes at 4°C. The supernatant was discarded. The pellet was re-suspended in 2ml of reagent B (Nucleon kit). 500µl of sodium perchlorate (Nucleon kit) were added and the solution mixed by gently inverting. 2ml of chloroform were added. This solution was mixed by gently inverting and then spun for 3 minutes at 3,600rpm. 300µl Nucleon resin were added and the tube was re-spun at 3,600rpm for 3 minutes. The clear upper phase was transferred to a new tube using a pipette. 2x the volume of cold absolute ethanol (approximately 5ml) were added and the tube inverted several times until the DNA was precipitated. The DNA was pelleted at 3,600rpm for 5 minutes. The supernatant was discarded and the pellet washed in 2ml cold 70% ethanol and re-centrifuged as before. The supernatant was discarded and the pellet air-dried for 10 minutes. The DNA was re-suspended in 50µl dH_2O.

2.1.3 DNA extraction from whole blood

3-10 ml of blood were transferred into a 50ml polypropylene tube. 4x the volume of reagent A (Nucleon kit) was added and the solution was mixed for 4 minutes in a rotary mixer. The solution was then centrifuged for four minutes at 1,000rpm. The supernatant was discarded without disturbing the pellet. 2ml reagent B was added and

the protocol proceeded as for DNA extraction from cultured cells as detailed above (2.1.2).

2.1.4 Preparation of agar plates.

LB agar contents are listed in section 2.8. The pH was adjusted to 7.0 using HCl. Following autoclaving, the agar was allowed to cool to approximately 55°C. 1 ml of antibiotic, chloramphenicol (25µg/µl) or kanaymcin (20µg/µl) was added. The agar was poured into petri dishes and stored at 4°C.

2.1.5 Bactierial artificial chromosome and P1 artificial chromosome culture.

Genomic probes, Bacterial Artificial Chromosomes (BACs) and P1 Artifical chromomsomes (PACs) were identified from the National Centre for Biotechnology Information (NCBI; www.ncbi.nim.nih.gov). These were obtained from the Wellcome Trust Sanger Institute as agar stabs. Clones were streaked on to agar plates, containing the appropriate antibiotic, and left overnight at 37°C. BAC clones from library RPCI-11 were chloramphenicol resistant. PAC clones from libraries RPCI-1, 3 and 5 were resistant to kanaymcin. Individual colonies were picked using a sterile pipette tip and transferred to 10ml LB broth (contents listed in section 2.8) containing 10µl of antibiotic, and left in a 37°C incubator, shaking overnight.

2.1.6 BAC and PAC DNA extraction

DNA was extracted using Qiagen Spin Miniprep kits and protocols (Qiagen, Crawley, UK). Following overnight liquid culture cells were pelleted at 5,000rpm for 5 minutes. The supernatant was discarded. The pellet was re-suspended in 250µl of P1 suspension buffer (Qiagen kit), containing RNase A, and transferred to a microfuge tube. 250µl P2 lysis buffer (Qiagen kit) were added and the tube inverted carefully 3-4 times. 350µl N3 neutralisation buffer (Qiagen kit) were added and the tube inverted gently 3-4 times and centrifuged for 10 minutes at 13,200rpm. The supernatant was decanted into a QIAprep column (Qiagen kit) and centrifuged for 1 minute at 13,200rpm. The flow-through was discarded. The spin columns were washed with 0.5ml of PB buffer (Qiagen kit) and centrifuged for 1 minute at 13,200rpm. A second

washing using 0.75 ml PE buffer (Qiagen kit) was carried out and centrifuged for 1 minute at 13,200rpm. The flow through was discarded and the column re-spun for 1 minute to remove any residual wash buffer. DNA was eluted into a 1.5ml microfuge tube using 50µl dH_2O, added to the centre of each column and left standing for a minute. The tubes were then centrifuged for 1 minute at 13,200rpm and stored at -20°C.

2.1.7 PCR amplification.

Amplification was carried out in a 15µl reaction mix of 10% 1 x DNA polymerase buffer, 10% 2.5µM dNTPs, 10% 2.5µM forward primer, 10% 2.5µM reverse primer, 50-100ng DNA, 0.2µl *Taq* 1u/µl DNA polymerase, and made up to 15µl with dH_2O. PCR cycle conditions: initial denaturation for 4 minutes at 94°C, followed by thirty-two cycles of 1 minute at 94°C, 52°C – 64°C depending on primer T_m for 1 minute, 72 °C for 1 minute and a final extension at 72 °C for 10 minute. PCR was carried out using an MJ research thermocycler.

2.1.8 Agarose gel electrophoresis.

PCR products were run out on a 1-2% ethidium bromide agarose gel with 1 x Ficoll Orange loading buffer. DNA was visualised using a Molecular Dynamics fluorimager 595 or UV-illuminometer.

2.2 Microsatellite analysis.

2.2.1 Identification and design of PCR primers

PCR primers amplifying polymorphic microsatellite markers were identified using the NCBI database (http://www.ncbi.nlm.nih.gov/). Where possible, published markers were used. However in regions of low density, additional polymorphic markers were identified from genomic clones mapping to the region. Primers were designed flanking dinucleotide repeats larger than 20bp in length using DNA star primer select software (DNASTAR Inc, Madison, US). Identified primers were put into a BLAST search (http://www.ncbi.nlm.nih.gov/) to ensure that they did not map to multiple loci. Primers were obtained from Qiagen Operon with a fluorescent label, FAM, TET or

HEX, attached to the 5' terminus of the forward primer to allow detection following gel electrophoresis. Optimisation of primer annealing temperatures was carried out on control DNA (protocol found in section 2.1.7). Primer sequence, modification and optimum PCR temperature can be found in appendix i.

2.2.2 Sizing of PCR products using an ABI 377 sequencer

PCR was carried out as outlined above. PCR products were sized using an ABI 377 DNA sequencer (Perkin-Elmer Life Sciences, Boston, US) and Genescan software (ABI, Fostercity, US). Multiplexing of primers in each lane was possible with 2µl of each primer PCR product of each patient DNA sample being pooled and made up to 20µl with dH_2O. Typically six primers could be pooled without PCR products overlapping. Similar sized fragments could be run in the same lane providing they were labelled with a different fluorescent dye. 0.5µl of pooled PCR sample was added to 2µl mix of 4 parts deionised formamide, 1 part Tamra size standard and 1 part Agarose loading buffer. The samples were denatured at 94°C for 3 minutes, and placed on ice to prevent re-annealing. Gel constituents for microsatellite analysis were 15ml 5% Gene-page plus, 150µl 10% APS, 15µl TEMED. 2µl of samples were loaded into each well and electrophoresed for 80 minutes. Data was extracted from the gel and analysed using ABI Prism Genescan software.

2.3 Fluorescent *in situ* hybridisation

2.3.1 Preparation of metaphase chromosomes from cultured cells.

100µl Colcemid (10µg/ml) were added to a growing cell culture and the solution was left at 37°C for 20 minutes. The cells were pelleted by centrifugation at 1,500rpm for 7 minutes. The pellet was re-suspended in 8ml KCl (0.56%), pre-warmed to 37°C, and left for 10 minutes. 2ml of methanol fixative (3 parts methanol: 1 part acetic acid) were added and the solution left for 30 minutes at 4°C. The solution was spun at 1,500rpm for 5 minutes and the pellet re-suspended in 10ml methanol fixative. The pellet was washed twice more in methanol fixative. The pellet was re-suspended in 0.5-1ml methanol fixative and several drops were pipetted on to clean glass slides. The slides were left at room temperature over night and placed on a hot plate for 1-2

hrs at 80°C the following day. The slides were incubated in a pepsin (0.005%) / HCl (0.01M) mix for 30min at 37°C. The slides were washed in 1x PBS for 5 minutes at room temperature, followed by 3 minutes washes in 70%, 90%, 100% ethanol. The slides were left at 37°C until dry.

2.3.2 Labelling of BAC and PAC DNA by nick translation

Protocols for the identification, propagation and DNA extraction of genomic clones can be found in section 2.1. DNA was directly labelled using nick translation. On ice, 1µg BAC/PAC DNA was labelled using 2.5µl 1x DNA Polymerase buffer, 2.5µl Nick Translation dNTP mix (see section 2.8 for contents), 1.75µl FITC (0.2mM), 5ml beta-mercaptoethanol (0.1M), 3µl Nick Translation enzyme mix (see section 2.8 for contents). The reaction mix was placed at 15°C for two hours and left on ice until used. Genomic probes were labelled with FITC. Chromosomal reference markers were labelled with Texas Red dye by following the same protocol.

2.3.3 FISH hybridisation and analysis

15µl of labelled DNA was precipitated in 3x cold ethanol with 5µg Cot1 DNA, 3µg sheared herring sperm and $^{1}/_{10}$ volume NaAc (pH 5.2; 3M) at -80°C for 15 minutes. Precipitated DNA was pelleted by centrifugation for 15 minutes at 14,000 rpm. The pellet was re-suspended in 15µl of FISH Hybridisation solution (see section 2.7 for contents). The DNA was hybridised to the metaphase slides. A coverslip was added and sealed with rubber cement. The slides were then placed on an 80°C heat block for 3 minutes to denature the metaphase chromosomes and the DNA. The slides were left overnight at 37°C. The slides were then washed three times at 60°C in 0.1xSSC for 5 minutes. Ensuring the slides didn't dry out the metaphases were counter-stained with Vectashield containg DAPI. Slides were viewed and analysed using a Zeiss epifluorescence microscope and Macprobe software.

2.4 Microarray-comparative genomic hybridisation

2.4.1 Preparation of the arrays

2.4.1.1 Identification of the target clones

Target BAC and PAC genomic clones were identified from the NCBI database (http://www.ncbi.nlm.nih.gov). These were obtained from the Wellcome Trust Sanger Institute as agar stabs. DNA was extracted from inoculated cultures using standard miniprep methodologies (sections 2.1.5 and 2.1.6). Genomic DNA was FISH-mapped to metaphase chromosomes (section 2.3).

2.4.1.2 Testing BAC/PAC clones for phage contamination.

Agar plates were prepared as before with no antibiotic and placed into a 37°C incubator for at least an hour. A DH10B *E.coli* culture was grown up over night in 10ml LB broth at 37°C. 4g agarose were added to 500ml LB broth, microwaved until the agarose had dissolved and placed in a 45°C waterbath. 10ml of *E. coli* culture were added to the agarose/LB broth and mixed well. 10 ml of broth/culture mix were poured on to the incubated agar plates. Once the plates had set, genomic clones were streaked out using a sterile tip and placed in a 37°C incubator overnight. Plates were checked the following day to see if the bacterial lawn had been lysed. Phage contamination results in lysis of the bacterial lawn and would appear as translucent areas on the bacterial lawn.

2.4.1.3 Amplification of BAC and PAC DNA using DOP-PCR.

Degenerate oligonucleotide primer-PCR was carried on genomic DNA extracted from BACs and PACs using three DOP-PCR primers in separate reactions. The design of the DOP-PCR primers is discussed in more detail in section 4.1. To ensure no genomic contamination DOP-PCR was carried out in a sterile environment. Amplification was carried out in a 50µl reaction mix of 5µl TAPS2 buffer, 5µl DOP primer (20µm), 4µl dNTPs (2.5mM), 2.5µl 1% W1, 0.5µl 5u/µl Taq polymerase, 28µl H_20, 5ng BAC/PAC DNA. PCR programme: 94°C for 3 minutes, 94°C for 1.3 minutes, 30°C for 2.30 minutes, ramp at 0.1°C per second to72°C, 72°C for 3 minutes, go to step 2 for 9 cycles, 94°C for 1 minutes, 62°C for 1.30 minutes, 72°C for 2

minutes, go to step 7 for 29 cycles 72°C for 8 minutes, 12°C for ever. 5µl of each product were run on a 2.5 % agarose gel and visualised using a UV-illuminometer.

2.4.1.4 Attachment of 5' amino group using amino-linked PCR primers.

Amino PCR was carried out to attach an amino group to the 5' terminus of the DOP-PCR product to facilitate attachment to the amine-binding glass slide. This was carried out using PCR primers with an amino-group at their 5' end (section 2.3.2). Amino-PCR was carried out on DOP-PCR amplified BAC and PAC DNA. Conditions were as follows: 6µl aminolinking buffer, 6µl dNTPs (2.5mM), 3µl aminoprimer (20µM), 0.6µl 5u/µl *Taq* polymerase, 42.4µl H_2O, 2µl DOP-PCR product. PCR programme using MJ thermocyclers: 95°C for 10 minutes, 35 cycles of 95°C for 1 minutes, 60°C for $1^1/_2$ minutes, 72°C for 7 minutes, 72°C for 10 minutes, held at 10°C.

2.4.1.5 Printing of the array-slides

40µl of the amino-linked products for each clone were combined and 39µl 4x Microarray Spotting Buffer (see section 2.8 for contents) were added to each. The arrays were printed onto coated amine-binding slides at the Wellcome Trust Sanger Institute array facility using a BioRobotics MicroGrid II Pro arrayer. Protocols for the slide printing can be found at http://www.sanger.ac.uk/Projects/Microarrays/. Each clone was printed in triplicate. Clones from the entire length of chromosome 1, with a resolution of 1 every Mb, as well as *drosophila* clones were obtained from collaborators at the Wellcome Trust Sanger Institute, details of which can be found at (http://www.ensembl.org/Homo_sapiens/cytoview). These were obtained already processed in spotting buffer and were also printed onto the slides in triplicate

2.4.2 Array hybridisation

2.4.2.1 Random-prime labelling of genomic DNA

0.3µg of DNA was labelled using an Invitrogen Bioprime labelling kit (Invitrogen Life Technologies, Paisley, UK). 0.3µg DNA were added to 40µl 2.5X Random Primers Solution (Invitrogen kit) and made up to 84µl with water. The DNA was denatured in a heat block for 10 minutes at 100°C, and immediately cooled on ice. On ice 10µl 10X dNTP mix (see section 2.8), 4µl Cy3 or Cy5 labelled dCTP (1mM), 2µl

Klenow fragment (Invitrogen kit) were added. The reaction was incubated at 37°C overnight. The reaction was stopped by adding 10µl stop buffer (Invitrogen kit). The labelled DNA was cleaned using Micro-spin G50 columns. 50µl of labelled DNA were added to the Micro-spin columns and spun for 2 minutes at 4,000rpm.

2.4.2.2 Precipitation of pre-/hybridisation solutions.

Pre-hybridisation solution: precipitate 40µl Herring Sperm (10mg/ml), 67.5µg human Cot1 DNA with 12µl NaAc (3M; pH5.2) and 300µl 100% cold ethanol. Hybrisiation solution: precipitate 90µl of each labelled DNA with 67.5µg human Cot1 DNA, 27µl NaAc (3M; pH5.2) and 700µl 100% cold ethanol. Both solutions were precipitated overnight at -20°C.

2.4.2.3 Resuspension of DNA

Precipitated DNA was spun for 15 minutes at 13,000rpm. The supernatant was discarded. 500µl 80% ethanol was added and re-spun at 13,000rpm for 2 minutes. Supernatant was removed and the samples were re-spun at 13,000rpm for 1 minute. The remaining supernatant was removed using a pipette. Labelled DNA was resuspended in 30µl pre-warmed (70°C) hybridisation solution and 3µl yeast tRNA (100µg/µl). DNA was denatured for 10 minutes at 70°C, before being transferred to 37°C for 60 minutes. Pre-hybridisation DNA was re-suspended in 70µl pre-warmed hybridisation buffer, and denatured at 70°C.

2.4.2.4 Microarray hybridisation

Two layers of rubber cement were applied around the array grids. Pre-hybridisation solution was applied to the well, and the slide transferred to a humidity chamber, containing tissue paper soaked in 2xSSC/40% formamide, and placed on a rocking table for 60 minutes. The pre-hybridisation solution was removed with a pipette and replaced with the hybridisation mix. Transferred to a humidity chamber, the chamber was sealed with para-film placed in a 37°C incubator for 48 hours.

2.4.2.5 Washing and analysis of array slides

Rubber cement was removed from the slide and the slide was washed in PBS/0.05% Tween 20. The slide was transferred to fresh PBS/0.05 Tween 20 solution for 10

minutes, after which it was placed in pre-heated (42°C) 50% formamide/ 2x SSC solution. The slides were rewashed in PBS/0.05 Tween 20 for 10 minutes and air dried. Slides were scanned using an Axon 4000b Scanner (Axon Instruments, Burlingham, US). The acquired images were quantified using Genepix 3.0 software (Axon Instruments). Genepix 3.0 software calculated the fluorescent intensities of all spots once the local background had been subtracted. To correct for non-specific hybridisation to spotted DNA the mean intensity of all the *drosophila* clones was subtracted for each fluorochrome from each of the human clones before ratio calculation. Spots with fluorescence intensities less than the mean *drosophila* dye intensities were rejected from further analysis. The test-over-reference ratio was calculated for each spot and normalised by dividing against the mean ratio of all the human clones. Ratios were accepted if the normalised ratios were within 20% of the median normalised ratio for the clone triplicate. If one result was outside that range, it was discarded from further analysis. If two were outside, all three were discarded from analysis. The mean ratio was calculated for each of the accepted triplicate ratios.

2.5 Reverse transcription-PCR

2.5.1 RNA extraction from embryonic tissue

The collection and use of human embryonic and fetal material was carried out following ethical approval from the Southampton and South West Hampshire Joint Local Research Ethics Committee. Written consent for the use of embryos was obtained from women undergoing termination of pregnancy. Human embryos were collected following medical or surgical termination. Fetal dissection was undertaken by the Developmental Genetics Group, Division of Human Genetics, University of Southampton. Human embryonic heart and tissue sections were homogenised in TRI REAGENT. 1ml of reagent was used for 50-100mg of tissue. The samples were allowed to stand for 5 minutes at room temperature. 0.2ml chloroform per ml of TRI REAGENT were added. The sample was shaken vigorously for 15 seconds and allowed to stand for 10 minutes. The tube was centrifuged for 15 minutes at 10,000 rpm at 4°C. The colourless upper aqueous phase, containing the RNA, was transferred into a fresh tube. 0.5ml isopropanol per ml of TRI REAGENT was added to the

sample and mixed. The sample was left standing for 10 minutes at room temperature followed by centrifugation at 10,000rpm for 10 minutes at 4°C to pellet the RNA precipitate. The supernatant was discarded and the pellet washed by adding 1ml 75% ethanol. The sample was then vortexed and centrifuged at 8,000rpm for 5 minutes at 4°C. The pellet was air-dried for 10min and re-suspended in DEPC H_2O.

2.5.2 DNase treatment of RNA samples

DNase was used to remove any residual genomic DNA. The following DNase digestion reaction was set up. 4µg RNA, 1µl RQ1 RNase-free DNase 10x buffer, 1µg/µ1 of RNA RQ1 RNase-free DNase, 1µl nuclease-free water was added to a final volume of 10µl. The reaction was incubated at 37°C for 30 minutes. 1µl RQ1 DNase stop solution was added. The DNase was inactivated by incubation for 10 minutes at 65°C. Following DNase treatment RNA was stored at -80 °C

2.5.3 First strand cDNA synthesis.

mRNA was converted to cDNA using an oligo (dt) primer which binds to the mRNA 3' poly (A) tail. Initial reaction containing 4µg RNA, 5µl dNTPs (0.5mM), 0.5µg Oligo dT primer, 4µl DEPC H_2O was placed in an MJ research thermocycler for 5 minutes at 65°C. The following reagents were then added: 4µl 5x cDNA buffer solution (section 2.8 for contents), 1µl RNase inhibitor, 2µl DTT (0.1mM). This was placed at 42°C for 2 minutes. 200U of Superscript Reverse Transcriptase was added. The reaction was carried out at 42°C for a further 50 minutes, 70°C for 15 minutes. cDNA was stored at -20°C

2.5.4 PCR amplification from cDNA.

Gene specific primers were designed from sequence available at the ensemble database (http://www.ensembl.org) using DNA Star software. Primers were designed to be intron spanning to ensure that any genomic contamination would not give a false positive result. PCR amplification and product visualisation was carried out as detailed in sections 2.1.7 and 2.1.8.

2.6 Immunohistochemistry analysis

2.6.1 Preparation of tissue sections

Preparation of tissue sections was carried out by the Developmental Genetics Group, Division of Human Genetics, University of Southampton

2.6.1.1 Tissue fixation and paraffin-embedding of fetal sections

Tissue was placed in 10 – 50ml of chosen fixative (4% paraformaldehyde (PFA) in PBS or Methanol acetone:acetone:Depc H_2O in a ratio of 2:2:1) and left to rock overnight. Fixative was removed and 70% ethanol (made with DEPC H_2O) was added. This was placed on a rocker for two hours. This was repeated with 80%, 90%, 100% ethanol and 100% chloroform. The tissue was left rocking in chloroform overnight. The tissue was placed in a new tube and filled with paraffin wax and left for 2 hours at 72°C. The paraffin wax was replaced with fresh wax and left for 2 hours at 72 °C. This step was repeated. The tissue was placed in a metal cassette and left on a cold plate to set for 30 minutes. This was stored at 4 °C until used.

2.6.1.2 Cutting and mounting of tissue

Paraffin blocks were mounted onto a microtome (Leica RM 2135, Leica Instruments, Germany). The microtome and all tools were cleaned with DEPC ethanol. Strips of embedded tissue were cut at 5µm intervals and placed onto RNase-free TESSPA coated glass slides. Four sections were placed on each slide. Every ninth section was put onto a separate slide for heamatoxylin and eosin staining. 0.1% DEPC-treated H_2O was squirted under each section and the slides were heated to 42 °C until the wax had melted. The water was aspirated and the slides transferred to 37 °C overnight. The slides were stored at 4 °C until used.

2.6.1.3 Heamatoxylin and eosin staining

Every ninth section cut from the paraffin block underwent haematoxylin and eosin staining. Paraffin was removed from the slides through immersion in xylene for 5 minures, followed by rehydration for 3 minutes each in 100% ethanol, 70% ethanol and H_2O. Slides were stained in Harris haematoxylin for 2-3 minutes, rinsed in H_2O and immersed in acid alcohol and bluing solution for 10 seconds each. Counter-stain

with eosin for 30-40 seconds was followed by graded dehydration in 50%, 70% ethanol, and 3 minutes in 100% ethanol. Slides were placed in xylene for 5 minutes and mounted in entellan

2.6.2 Protein detection using horseradish peroxidase/3,3' diaminobenezidine.

Paraffin-embedded sections were rehyrated by immersion in xylene for 5 minutes, 100% ethanol for 3 minutes, 70% ethanol for 3 minutes and dH_2O for 3 minutes. To remove endodgenous peroxide slides were in left a solution of 250ml PBS (pH 7.4) and 7.5ml hydrogen peroxide for 20 minutes. Antigen unmasking was carried out using trypsin (1mg/ml) and immersion in boiling sodium citrate (10mM). Tissue sections were incubated at 4°C overnight with primary antibody in solution with animal serum and immunobuffer in a humidified chamber. The slides were washed in PBS three times for 5 minutes. Secondary antibody was applied to each section for two hours at 4°C in a humidified chamber. Slides were washed in PBS as before. Streptavidin-HRP was applied to each section for an hour at 4°C in a humidified chamber. Slides were washed in PBS as before. DAB stock solution was placed on the sections for 10 minutes. Excess solution was then removed. DAB working solution was applied to each section for 3 minutes. The slides were washed in PBS as before. Sections were then stained with toludine blue for 2 minutes followed by washing in distilled water for 2minutes. The slides were put through an ethanol series of 10 seconds 70% ethanol, 10 seconds 90% ethanol and 2 minutes at 100% ethanol. Slides were twice immersed in xylene for 2 minutes and were then mounted in entellan. Sections were viewed using a Zeiss Axiovision imaging system.

2.7 Mutation screening

Mutation screening was performed by denaturing high performance liquid chromatography analysis (dHPLC) and by direct sequencing. Primers were designed to amplify each exon with approximately 50bp 5' flanking intronic sequence. Primers were designed from sequence available at http://www.ensembl.org using DNA star software.

2.7.1 dHPLC analysis

Each exon was amplified from patient DNA by PCR as detailed in section 2.1.7. Heteroduplexes were obtained by denaturing the DNA at 94°C. For each exon, dHPLC analysis was performed on a Transgenomic Wave DNA Fragment Analysis System using a DNAsep column (Transgenomic, Crewe, UK). Column temperatures were calculated based on the sequence and length in bp of the amplicon using Wavemaker software (Transgenomic). Column temperatures used for each exon can be found in appendix i. 10µl of the PCR mixture was injected into the column The separation conditions for analysis were a flow rate of 0.9ml/min and the ratio of buffers A and B was adjusted to elute the product between 3.5 and 6 minutes. DNA was detected by monitoring the absorbance at 260nm. Heteroduplex peaks were investigated by direct sequencing analysis.

2.7.2 Direct DNA sequencing.

PCR amplification was carried out in 30µl reactions. Products were run on a 2% agarose gel with 1.5µl of 10x Ficoll orange loading dye and visualised using a UV-illuminometer. Gel extraction of PCR products was carried out using a Qiagen kit and protocol (Qiagen, Crawley, UK). PCR products were cut from the 2% agarose gel using a scalpel, and placed in 1.5ml eppendorf. 500µl buffer QG (provided in the Qiagen kit) were added to each eppendorf which were then incubated in a 50°C water bath for 10 minutes. The eppendorfs were vortexed to help the agarose to dissolve. The contents were transferred to QIAquick spin column. The column was centrifuged for 1 minute at 13,000 rpm. The flow-through was discarded. 500µl QG were added to the column to remove any remaining traces of agarose, and centrifuged as before. The flow-through was discarded and 750µl of Buffer PE was added to the column. The column was left to stand for 3 minutes before centrifuging for 1 minute at 13,000 rpm. The flow-through was discarded and the column spun for an additional minute to remove any residual ethanol. The columns were placed in a new eppendorf. 30µl dH_2O were added to the centre of the column, left to stand for 1 minute and centrifuged for a minute to elute the DNA.

Sequencing reactions were carried out using the ABI PRISM dye Terminator Cycle Sequencing Ready reaction kit (Perkin Elmer Life Sciences, Boston, US). The reactions consisted of 8µl terminator ready reaction mix, 3-10ng DNA PCR product, 3.2pmol forward or reverse primer and made up to a volume of 20µl with dH_2O. Initial denaturation was carried out at 96°C for 5 minutes. This was followed by 25 cycles of 96°C for 10 seconds, 50°C for seconds, and 60°C for 4 minutes. The products were held at 4°C until used. The products were isopropanol precipitated with 80µl of 75% isopropanol for 15 minutes, followed by centrifugation for 20 minutes at 13,200rpm. The supernatant was aspirated using a syringe. 250µl of isopropanol were added to the tubes, and centrifuged at 13,200 for 5 minutes. The supernatants were removed using a syringe and the samples were dried in a thermo cycler for 1 minute at 90°C.

PCR pellets were re-suspended in 4µl of loading mix, containing 5 parts formamide to 1 part Agarose loading buffer. Products were denatured at 96°C for 3 minutes and maintained on ice. Products were electrophoresised using an ABI 377 DNA sequencer (Perkin Elmer Life Sciences) for $3^1/_2$ hours. Gel constituents consisted of 30ml 5% Gene-page plus, 300µl 10% APS, 30µl TEMED. Sequence analysis was carried out using Chromas and DNA Star software (DNASTAR Inc, Madison, US).

2.8 General laboratory solutions and buffers

Acid Alcohol:	95%ethanol, 1%HCl
Aminolinking buffer:	500mM KCl, 25mM $MgCl_2$, 50mM Tris pH8.5
5 x cDNA buffer solution:	250mM Tris-HCl pH 8.3, 375 mM KCl, 15mM $MgCl_2$
DHPLC buffers:	All buffers were made using HPLC grade water
buffer A:	0.1M Triethylammonium acetate
buffer B:	0.1M TEAA, 25% acetonitrile
buffer C:	75% acetonitrile
buffer D	25% acetonitrile
1 x DNA polymerase buffer:	50mM KCl, 1.5mM $MgCl_2$, 10mM Tris-HCl pH9, 0.1% triton
10 x Ficoll gel loading buffer:	0.25 % Orange G, 25 % Ficoll, 0.25 M EDTA.
FISH hybridisation solution:	50% deionised formamide, 40% dextran sulphate and 10% 20x SSC
Immunohistochemistry buffer:	PBS, 0.1% Triton X100.
LB Broth pH 7.0:	10 g NaCl, 10 g tryptone, 5 g yeast extract, dH_2O to 1 L, autoclave.
LB agar:	Add 20 g bacto agar to LB broth.
Methanol Fixative:	3 methanol:1 acetic acid.
Microarray hybrisisation buffer:	50% formamide, 10% dextran sulphate, 0.1% Tween 20, 2x SSC, 10mM Tris pH 7.4.
4x Microarray spotting buffer:	1M sodium phasphate buffer, pH8.5, 0.001 sarkosyl.
Nick translation dNTP mix:	0.5mM dATP, dGTP, dCTP, 0.4mM dTTP.
Nick translation enzyme mix:	100µl glycerol, 20µl DNase buffer, 22µl 5u/µl DNA polymerase, 1µl 1u/µl DNAase, 66µl dH_2O

Random prime labelling dNTP mix:	0.5 mM dCTP, 2 mM dATP, 2mM dGTP, 2mM dTTP in TE buffer
20 x SSC:	3 M NaCl, 0.3 M Tri-sodium citrate
10 x TAE pH 8.2:	40 mM Trisma Base, 2 mM 0.5 M EDTA, 20 mM sodium acetate.
TAPS2 buffer:	250mM TAPS pH 9.3, 166mM $(NH_4)_2SO_4$, 25mM $MgCl_2$ stored at -20°C. 33µl BSA and 7µl β-mercaptoethanol were added to the buffer just prior to use.
10 x TBE:	540 g Trisma base, 275 g Orthoboric acid, 200 ml 0.5 M EDTA, dissolve in dH_2O to 5 L.
TE Buffer:	10 mM Tris-HCl pH 7.4-8.0, 1 mM EDTA pH 8.0, autoclave.

2.9 Materials

Plasticware was obtained from Fischer, Loughborough, UK.

Acetic acid	-	Fischer, Loughborough, UK
Acetonitrile	-	Sigma-Aldrich, Poole, UK
Agar	-	Oxoid, Bsingstoke, UK
Agarose	-	Sigma-Aldrich
Agarose loading buffer	-	Applied Biosystems, Warrington, UK
APS	-	Sigma-Aldrich
Beta-mercaptoethanol	-	BDH, Poole, UK
BSA	-	Sigma-Aldrich
Chloramphenicol	-	Sigma-Aldrich
Chloroform	-	Fischer
Colemid	-	Invitrogen Life Technologies, Paisley, UK
Cot1 DNA	-	Roche, Welwyn Garden City, UK
Cy3/Cy5	-	Perkin-Elmer, Beaconsfield, UK
DAB	-	Fischer
DAPI	-	Vector laboratories, Peterborough, UK
DEPC	-	Sigma-Aldrich
Dextran sulphate	-	Sigma-Aldrich
DNase	-	Promega, Southampton, UK
DNase stop solution	-	Promega
dNTPs	-	Roche diagnostics
EDTA	-	Sigma-Aldrich
Entellan	-	Merck, Hoddesdon, UK
Eosin	-	Thermoshandon, Runcorn, UK
Ethidium Bromide	-	Sigma-Aldrich
Ficoll	-	Sigma-Aldrich
FITC	-	Vysis, Maidenhead, UK
Formamide	-	Sigma-Aldrich
Genepage plus	-	Ameresco, Solon, Ohio, US
Glycerol	-	BDH
Glass slides	-	BDH
Glass slides (amine binding)	-	Motorola Life Sciences, California, US
haematoxylin	-	Thermoshandon
HCl	-	BDH
Herring sperm sperm	-	Sigma-Aldrich
Hydrogen peroxide	-	Sigma-Aldrich
Isopropanol	-	Fischer
Kanamycin	-	Sigma-Aldrich
KCl	-	BDH
Methanol	-	Fischer
Micro-spin G50 columns	-	Amersham, Little Chalfont, UK
NaCl	-	Sigma-Aldrich
NaAc	-	BDH
Orange G	-	Sigma-Aldrich
Orthoboric acid	-	Sigma-Aldrich

Oligo dT primer	-	Invitrogen Life Technologies
PBS	-	Sigma-Aldrich
Pepsin	-	Sigma-Aldrich
RNasin	-	Promega
RPMI	-	PAA laboratories, Yeovil, UK
RNase-free DNase	-	Promega
RNase-free DNase 10x buffer	-	Promega
DNase stop solution	-	Promega
Salmon sperm DNA	-	Sigma-Aldrich
Sodium Citrate	-	Sigma-Aldrich
Superscript RT	-	Invitrogen Life Technologies
Tamra size standard	-	Applied Biosystems, Warrington, UK
TAPS	-	Sigma-Aldrich
Taq DNA polymerse	-	Promega and Perkin-Elmer
TEAA	-	Transgenomic, Crewe, UK
TEMED	-	Sigma-Aldrich
Texas red	-	Vysis
Toludine Blue	-	Sigma-Aldrich
Trinitron x100	-	Sigma-Aldrich
Trisma Base	-	Sigma-Aldrich
Tris-HCl	-	Sigma-Aldrich
TRI Reagent	-	Sigma-Aldrich
tRNA	-	Invitrogen
Trypsin	-	Sigma-Aldrich
Tryptone	-	Oxoid
Tween 20	-	Sigma-Aldrich
Vectashield (containing DAPI)	-	Vector Laboratories
Yeast extract	-	Oxoid
Xylene	-	Fisher

Chapter 3: Defining a critical 1q42-44 disease region

3.1 Introduction

3.1.1 Deletions within 1q42-44

Microscopically visible telomeric deletions can cause specific malformation and mental retardation syndromes such as 3p- (OMIM 606217), 4p- (Wolf-Hirschorn syndrome; OMIM 194190), 5p- (cri du chat syndrome; OMIM 123450), 9p-, 13q-, and 18p- syndrome. These can often be recognised through a series of characteristic phenotypes. Distal deletion of 1q (q42 or q43-ter) gives rise to a consistent pattern of malformations. Since the first report by Mankinen in 1976, over 30 cases with microscopically visible 1q deletions have been described (Mankinen et al., 1976). The phenotypic similarities seen between patients have lead to attempts to define a specific 1q deletion syndrome. Juberg first postulated the possibility of a new deletion syndrome within 1q43 from the similarity of a male infant with monosomy 1q43, resulting from a translocation between chromosomes 1 and 16, to five previous case reports (Juberg et al., 1981). No significance was attached to the fact that the patient was trisomic for 16q24-ter. Several subsequent reports supported the definition of a specific 1q42-ter syndrome (Johnson et al., 1985; Manouvrier-Hanu et al., 1986). However, Watson suggested that, although there were phenotypic similarities between patients, these are seen in other cases of chromosomal abnormality, and that no single anomaly or pattern of malformations occurred consistently enough to permit a specific diagnosis (Watson et al., 1986). Meinecke and Vogtel following characterisation of two patients and a comparison with published cases concluded that terminal 1q deletion is a definable and recognisable syndrome, but agreed with Watson that the individual manifestations observed are common to many chromosomal disorders (Meinecke and Vogtel, 1987). A possible mechanism for the deletion is the presence of a fragile site within 1q42.1 (Pelliccia et al., 1998).

Patients present with an array of developmental disorders. Central nervous defects significantly associated with loss of 1q42-44 include microcephaly, hydrocephalus and agenesis of the corpus callosum, the latter being a highly significant association (Brewer et al., 1998). Some may also suffer from developmental delay, seizures and microcephaly. Genital defects, such as hypospadias, have also been documented in these patients (Brewer et al., 1998). Craniofacial defects are common and can include

a full round face with a prominent forehead, upward slanting palpebral fissures, epicanthic folds, a short, broad nose with a flat nasal bridge, thin lips with down-turned corners of the mouth, micrognathia, low set ears, an abnormal palate and a short neck.

Cardiorespiratory complications are cited as the prominent cause of death. The cardiac anomalies vary, ranging from ventricular septal defects, to life-threatening defects such as hypoplastic left heart (see table 3.1). Monosomy of 1q42-44 is significantly associated with ventricular septal defects (Brewer et al., 1998). Several authors identified a more proximal interstitial deletion extending from 1q32-42 (Al-Awadi et al., 1986; Sarda et al., 1992). Both had cardiac defects, as well as the characteristic facial appearance, agenesis of the corpus callosum, as well as malformations of the hands. Both infants died and no autopsy was performed. Youssoufian reported on a boy with a similar deletion who on birth was cyanotic and had no detectable heart rate (Youssoufian et al., 1988). Physical examination revealed low-set ears, mild webbing of the neck, and hypospadias. The child died at 8hrs of age due to respiratory failure. Although no cardiac defect was reported, autopsy was not performed to confirm normal anatomy. The majority of other case reports present with large terminal deletions extending from 1q42.3/ 1q43 to the telomere (Garani et al., 1988; Johnson et al., 1985; Kessel et al., 1978; Meinecke and Vogtel, 1987; Watson et al., 1986; Wright et al., 1986). Several cases of 1q monsomy giving rise to CHD result from unbalanced translocations (De Vries et al., 2001; Juberg et al., 1981; Merlob et al., 1989). The overlap seen between the proximal interstitial deletions of 1q32-42 and the large terminal deletions extending from 1q42.3 may suggest the presence of a CHD susceptibility locus within distal 1q42. Alternatively there may exist two non-overlapping CHD critical regions, one located proximally within 1q32-42 and the other distally within 1q42.3-q44. De Vries reported on two CHD patients who were shown to have deletions within 1q44, which may support the hypothesis that there are two critical cardiac regions. (De Vries, Knight *et al.* 2001). However, as Watson *et al* suggested, insufficient chromosomal resolution of many of the case reports may have lead to a lack of precision in assigning break sites (Watson et al., 1986).

Deleted region	Cardiac defect	Extracardiac defects	Reference
1q32-42	VSD, ASD	Craniofacial abnormalities (agenesis of the corpus callosum), renal hypoplasia, ectopic anus.	(Sarda et al., 1992)
1q32-42	Cardiac defect not specified	Craniofacial abnormalities (encephalocele, micrognathia), ectopic anus, weak cry, genital defects, a short neck	(Al-Awadi et al., 1986)
1q42-44	PDA, atresia of pulmonary valves, VSD with overriding aorta	Craniofacial dysmorphology including microcephaly and cleft palate, hypospadias	(Kessel et al., 1978)
1q42-44	VSD, valvular pulmonary stenosis	Facial dysmorphology, microcephaly, a short neck, genital and anal defects, ageneisis of the corpus callosum, psychomotor delay,	(Johnson et al., 1985)
1q42-44	HLH, atresia of the aortic arch, mitral and aortic valves, PDA	Micrcephaly, facial dysmorphology, partial absence of the corpus callosum, ocular defect	(Wright et al., 1986)
1q42.3 - 44	VSD	Reduced, speech and intellectual development, seizures, craniofacial dysmorphism, genital defects, short neck	(Meinecke and Vogtel, 1987)
1q42.3 - 44	VSD, patent foramen ovalve	Cleft palate, microcephalic, facial dysmorphology, retardation of motor, speech and intellectual development, abnormal digits	(Meinecke and Vogtel, 1987)
1q42.3-44	right ventricular hypertrophy	Severely growth retarded, genital abnormalities, facial abnormalities including a cleft palate, a short neck, abnormal digits	(Watson et al., 1986)
1q43-44*	Large VSD, and right overriding aorta	Microcephaly, facial dysmorphology, genital abnormalities, digit defects, weak cry	(Merlob et al., 1989)

Table 3.1: continued on the following page.

Deleted region	Cardiac defect	Extracardiac defects	Reference
1q43-44	VSD	Microcephaly, facial dysmorphology, partial agenesis of the corpus callosum, seizures, psychomotor delay,	(Garani et al., 1988)
1q43-44[1]*	Congenital heart defect not specified	Encephalocele, craniofacial defects, abnormal genetalia and anus, a short neck and digit defects	(Golabi et al., 1982)
1q43-44[2]*	VSD, increased pulmonary vascularity, biventricular hypertrophy	Genital abonormalities, facial dysmorphology, weak cry, microcephaly, convulsions, a short neck, defects of the hands	(Juberg et al., 1981)
1q44	VSD, subaortic stenosis	Hypospadias, kidney abnormalities, seizures, partial agenesis of the corpus callosum, developmental delay, charcteristic facial dysmorphology	(De Vries et al., 2001)
1q44*	Patent foramen ovalve	Cleft palate, hyposadias, severe microcephaly, characteristic facial dysmorphology, a short neck, poorly formed corpus callosum	(De Vries et al., 2001)

Table 3.1: Cardiac and extracardiac defects of 1q deletion case reports. Those cases marked with an *, the deletion results from inheriting a derivative chromosome 1, resulting from a balanced translocation from one of the parents. The remaining cases were *de novo*. VSD – ventricular septal defect; ASD – atrial septal defect; PDA – patent ductus arteriosus; HLH – hypoplastic left heart.

[1] A second pregnancy was diagnosed to have the same chromosomal abnormality by midtrimester amniocentesis. Pregnancy was terminated. Mother and maternal grandfather were shown to have the same balanced translocations.

[2] Previous history of a still born male with tetralogy of fallot (aortic transposition, pulmonary artery atresia, VSD, right ventricular hypertrophy). No chromosomal study was carried out.

Similar to other deletion syndromes, the cardiac defects show incomplete penetrance. Some individuals with overlapping deletions display no cardiac defect, but will present with similar craniofacial abnormalities, agenesis of the corpus callosum, characteristic facial dysmorphology, growth and development delay (Gentile et al., 2003; Manouvrier-Hanu et al., 1986).

3.1.2 Trisomy 1q42-1q44

Duplications of 1q42-44 have been reported. These most commonly occur as a result of an unbalanced re-arrangement, which may involve deletion of another chromosomal region. In such cases it can be difficult to determine which chromosomal imbalance is responsible for the malformations seen. Pure trisomy distal 1q has been reported infrequently, but has been suggested to give rise to a characteristic phenotypic syndrome. Seven cases have been published. Nowaczyk *et al.* recently reviewed these whilst presenting a case report of their own (Nowaczyk et al., 2003). Although one report gave little phenotypic data, similar abnormalities were identified between the patients. Several facial and urogenital abnormalities, such as hypospadias, were consistent with those seen in monosomy 1q42-44. Cardiac defects were involved in four out the seven cases (Clark et al., 1994; Duba et al., 1997; Nowaczyk et al., 2003; Steffensen et al., 1977). Cardiac defects included patent ductus artiosus, tricuspid valve dysplasia resulting in tricuspid regurgitation, VSD, pulmonary stenosis and aortic stenosis.

3.1.3 Disease genes mapping to 1q42-44

Several genes mapping to the 1q42-44 have been shown to have a role in human disease. Mutations in the *TBCE* gene were recently shown to be a cause of HRD/ Sanjad-Sakati syndrome (OMIM 241410) and Kenny-Caffey syndrome (Parvari et al., 2002). Both syndromes present with congenital hypoparathyroidism, mental retardation, microcephaly, facial dysmorphism and extreme growth failure. Kenny-Caffey Syndrome presents with the additional features of osteosclerosis and recurrent bacterial infections. Both are autosomal recessive disorders which were mapped to this region by linkage analysis (Diaz et al., 1998; Parvari et al., 1998). No cardiac defects were documented in these patients. A gene thought to be a cause of

schizophrenia, disrupted in schizophrenia 1 (*DISC1*), also maps to 1q42.1 (OMIM 605210). A balanced translocation (1;11)(q42.1;q14.3) was found to segregate with schizophrenia and related psychiatric disorders in a large Scottish family (St Clair et al., 1990). Cloning of the 1q42 breakpoint identified *DISC1* as the most likely gene (Millar et al., 2000). Although no schizophrenia has been documented in the setting of 1q42-44 deletions, neurological complications can be a feature (Murayama et al., 1991), possibly resulting from haploinsufficiency of *DISC1*. Mutations in the cardiac ryanodine receptor 2 (*RYR2*) are known to be a cause of arrthymogenic right ventricular dysplasia type 2 (Tiso et al., 2001). No congenital structural cardiac malformations were documented in these patients.

3.1.4 The molecular investigation of 1q42-44 CHD

The consistent presence of cardiac defects seen in association with chromosomal imbalance of 1q42-44 suggests that within this region is a critical cardiac disease gene; half dosasge of which is insufficient for normal development. At the time of writing no gene that maps to the 1q42-44 region has been implicated in the pathogenesis of CHD. The case reports of 1q42-44 monosomy present little, if any, molecular data, and no attempt has been made to define a critical 1q CHD susceptibility locus for prioritisation of candidate genes. The lack of molecular data within the literature makes it difficult to identify phenotype/genotype correlations and define adequate critical regions with which to target potential candidate genes. Molecular definition of a 1q42-44 CHD region was therefore a priority. Microsatellite analysis to identify non-inheritance of polymorphic alleles was carried out on a patient, patient 1, with a known cytogenetic deletion within 1q42-44 and her parents. Subsequent to the analysis of this patient, the DNA of ten further patients with known deletions and their parents were obtained. Microsatellite investigations were undertaken to ascertain whether there was overlap with the deletion of patient 1.

Most cases of suspected chromosomal abnormality seen in the clinic are referred for karyotyping analysis. The resolution of cytogenetic banding for detecting deletions is limited to approximately 3Mb. Submicroscopic deletions, those under 3Mb, may therefore not be detected by chromosomal banding. DiGeorge Syndrome was originally considered to be a rare syndrome until characterisation of the deletion

region and the introduction of fluorescent *in situ* hybridisation (FISH) as a routine diagnostic test. It is now known to be the most common deletion syndrome. However higher resolution techniques, such as FISH, are only requested if a specific abnormality is suspected. It is therefore possible that microdeletions within 1q42-44 are going undetected and that monosomy 1q42-44 is more common than currently thought. The finding of submicroscopic deletions would be a significant step forward in a deletion mapping strategy to identify candidate CHD genes. Having established a molecular CHD critical region it was therefore possible to screen for microdeletions using loss-of-heterozygosity analysis in a panel of CHD patients with normal karyotypes. Patients with a block of contiguous polymorphic markers amplifying single alleles were to be investigated further using FISH analysis to confirm the presence or absence of a deletion.

3.2 Aims

1. To define a molecular critical CHD region within 1q42-44 using microsatellite analysis to identify non-inheritance of parental alleles in patients with known 1q deletions.

2. To carry out loss-of-heterozygosity analysis and FISH analysis of a panel of CHD patients with normal karyotypes to identify submicroscopic deletions within the defined critical region.

Patient number	Cardiac defect
Patient 1	VSD, pulmonary stenosis
Patient 2	ASD
Patient 3	VSD, ASD
Patient 4	VSD, subaortic stenosis
Patient 5	ASD, multiple small VSD
Patient 6	VSD, PDA, pulmonary stenosis
Patient 7	Patent foramen ovalve
Patient 8	No cardiac defect
Patient 9	No cardiac defect
Patient 10	No cardiac defect
Patient 11	No cardiac defect

Table 3.2: The cardiac anatomy of patients 1-11 that were investigated using microsatellite analysis. VSD – ventricular septal defect; ASD – atrial septal defect; PDA – patent ductus arteriosus

3.3 Results

3.3.1 Molecular analysis of patient 1

Patient 1 presented with a ventricular septal defect (VSD), pulmonary stenosis, facial dysmorphology and an encephalcoele. Routine cytogenetic karyotyping was carried out and shown to be 46,XX,del(1)(q42.13q42.3). DNA was available from a peripheral blood sample and a lymphoblastoid cell line. Parental DNA was obtained from blood. DNA was extracted from cell-lines and blood following standard procedures detailed in section 2.1.

Following microsatellite analysis of patient and parental DNA the deletion size and position were determined. The patient was shown to have a *de novo* interstitial deletion. The child had inherited no paternal alleles. The proximal and distal boundaries are defined by markers DIS225 and D1S1149 respectively. This gives a maximal deletion size of approximately 9Mb. The results are summarised in table 3.3.

Locus	NCBI Position (Kb)	Genotype (bp)		
		Patient 1	Paternal	Maternal
D1S2880	216 925	130/132	130/132	126/130
D1S2871	218 544	232/234	216/232	216/234
D1S2763	220 768	165/167	165/167	165/167
D1S439	222 084	253/255	255/257	253/259
D1S1644	222 819	265/273	265/273	273
D1S1617	226 010	125/131	129/131	121/125
D1S225	227 466	219/221	221/233	219
D1S251	228 021	255	257/259	255/267
D1S2709	228 337	195	195	195
D1S459	228 801	182	184/196	180/182
B804 (RP4-804P4)	229 460	142	136/140	142/144
B895 (RP5-895L2)	230 123	281	283/285	281
D1S179	230 966	165	163/167	165
B781 (RP4-781K5)	231 047	242	230/240	240/242
D1S235	232 219	176	176/187	176/187
B940 (RP5-940F7)	232 415	319	319/324	319
D1S2850	232 956	147	147/149	147
B519 (RP11-182B22)	233 330	270	255/268	270/274
B44 (RP11-44E24)	233 694	190	190	190
B541 (RP11-541M23)	233 983	297	287/291	289/297
B193 (RP11-193H5)	234 387	292	292/294	286/292
B307 (RP11-307O1)	235 533	347	321/347	335/347
D1S517	235 686	199	214/217	199/214
B359 (RP11-359A17)	235 850	230	224/232	230/232
D1S1149	236 473	377/383	377	377/383
B177 (RP11-177F11)	236 840	236	230/236	236
D1S2785	237 202	166/168	166/174	168/182
D1S180	237 691	173/183	169/183	173

Table 3.3: Results of microsatellite analysis carried out on patient 1 and her parents. The location of markers and genomic clones are shown relative to the NCBI map (http://www.ncbi.nlm.nih.gov/; sequence as 10/04). Microsatellites prefixed with B- were designed from genomic clones, the name of which is bracketed, as detailed in section 2.2. The location given is that for the clone. Allele sizes were determined following electrophoresis in an AB1377 automated sequencer and analysis using genotyper software (see section 2.2). Patient 1 was shown to have a 9Mb interstitial deletion resulting from loss of the paternal alleles. Deleted markers and those suspected of being deleted are highlighted in red.

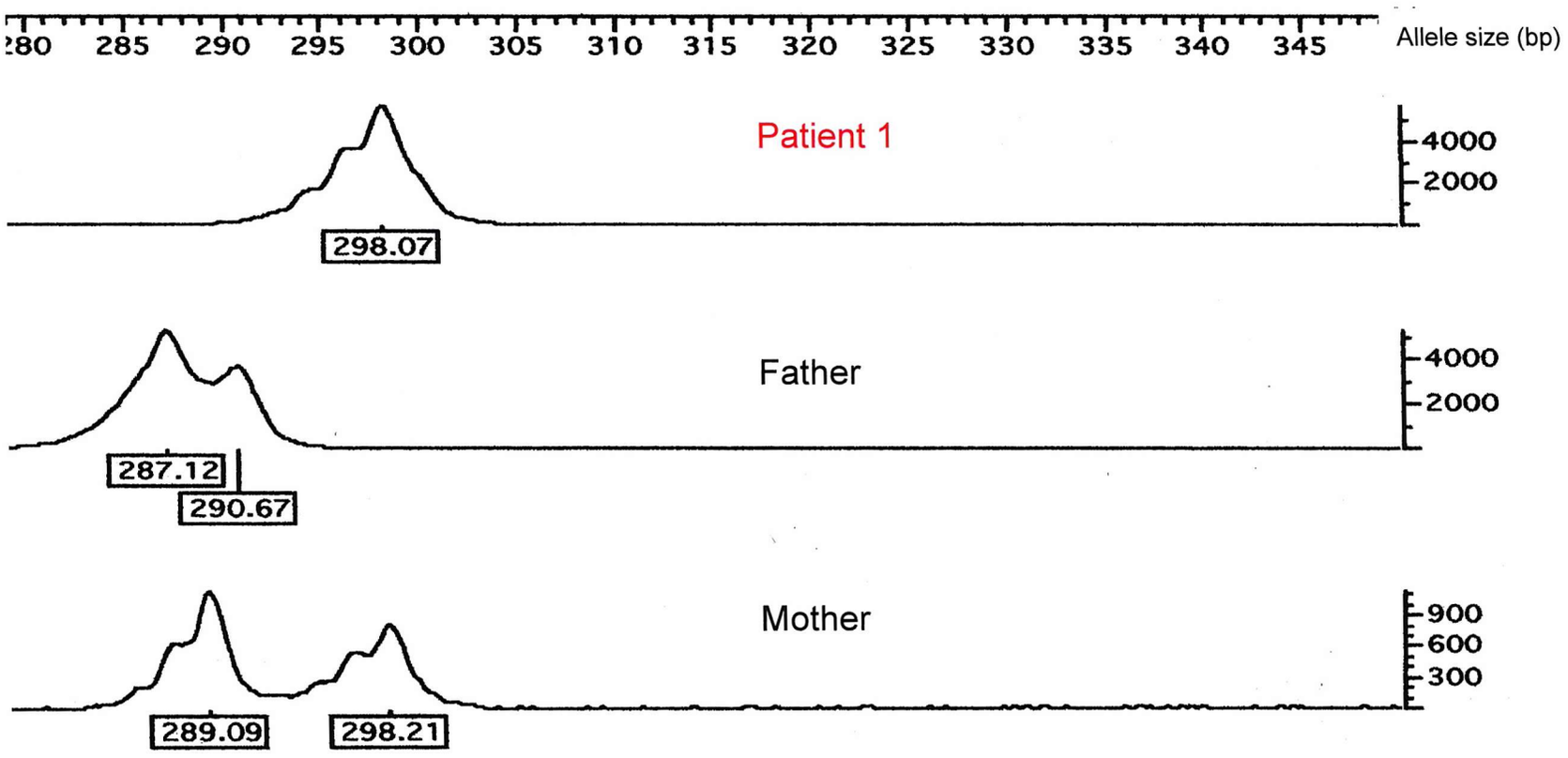

Fig 3.1: The result traces of microsatellite analysis carried out on patient 1 and parental DNA following amplification of microsatellite B541 (see table 3.3 for locus position). The number shown beneath each peak is the allele size in bp graded against the scale above. The figures to the right represent the intensity of the signal. The father and mother are both heterozygous at this locus. Patient 1 has inherited allele 298 from the mother, but no allele from the father confirming that the child was deleted for this marker.

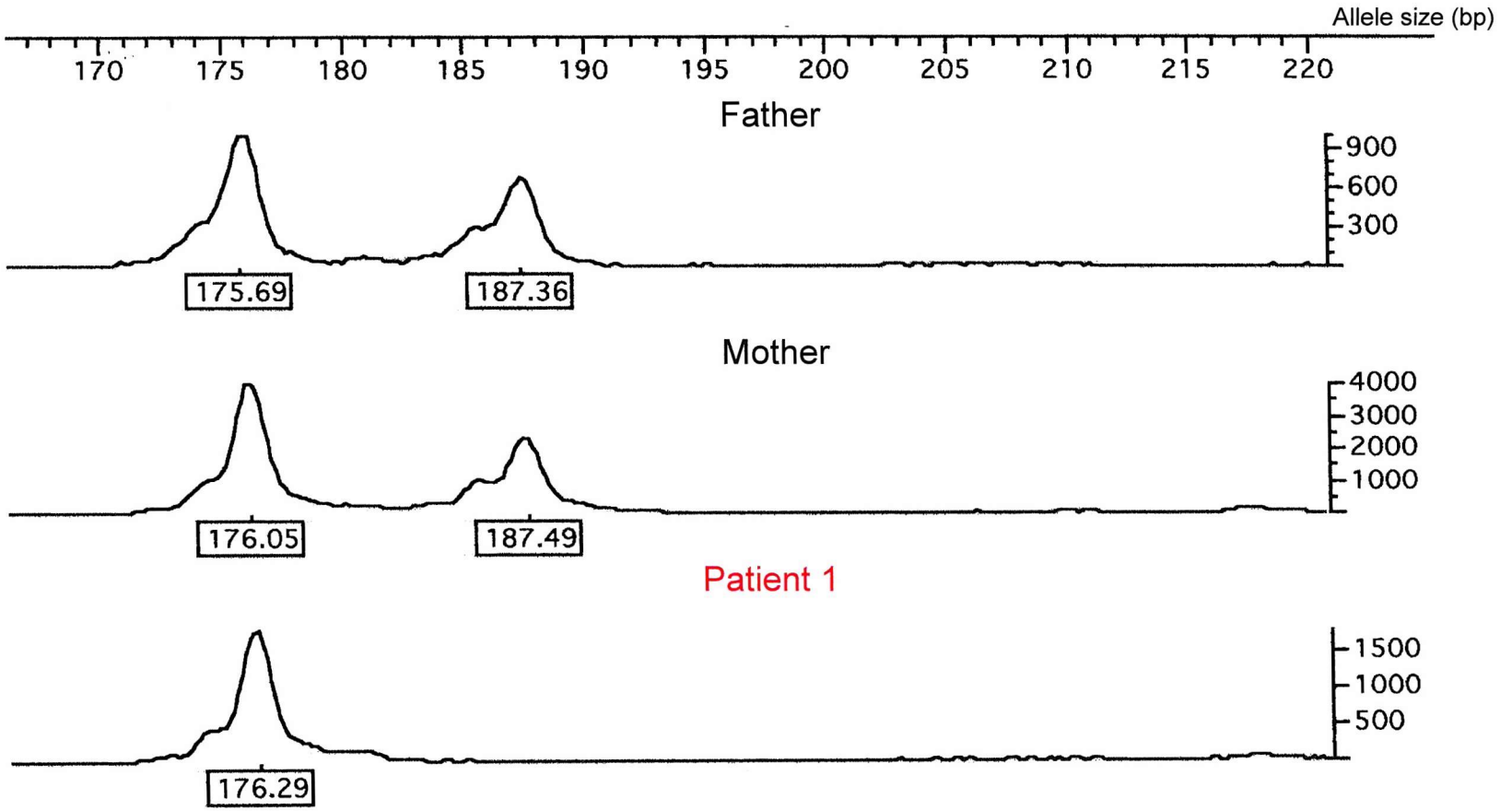

Fig 3.2: The result traces of microsatellite analysis carried out on patient 1 and parental DNA following amplification using microsatellite D1S235. The number shown beneath each peak is the allele size in bp graded against the scale above. The figures to the right represent the intensity of the signal. Both father and mother are heterozygous at this locus with the same allele sizes. The result for patient 1 is uninformative. It not possible to determine from this data whether patient 1 is homozygous, inheriting an allele from both parents, or is hemizygous inheriting only one allele. Given that this marker is located within the middle of the deletion identified with patient 1, it is assumed that the patient is deleted at this locus.

3.3.2 Molecular analysis of patients 2-11

Following analysis of patient 1, a further 10 patients were obtained who were known to have deletions within 1q42-44 region. Patient and parental DNA were available. Patients 2-7 had congenital heart disease. The nature of these defects can be found in table 3.2. Microsatellite analysis was carried out on all of these patients to ascertain if there was any overlap with the original critical region established from the studies of patient 1. Analysis was carried out as before. The results are shown in table 3.4. All patients had a terminal deletion except for patient 3 who was heterozygous for the distal marker D1S3739. The centromeric breakpoints varied. Patients 2, 3 and 4 all of whom had cardiac defects had deletions that overlapped with that of patient 1. Patient 2 had the largest deletion, 14 Mb, extending from marker D1S235 to the telomere. The remaining three CHD patients, 5-7, had deletions distal to the original critical region. Patient 7 had the smallest deletion extending from the telomere to D1S547 representing a deletion size of approximately 7Mb. There is over a megabase between the interstitial deletion of patient 1 and the terminal deletion of patient 7. The four non-CHD patients all had deletions distal to the patient 1 deletion.

Locus	NCBI Position	Patient 1 (CHD)	Patient 2 (CHD)	Patient 3 (CHD)	Patient 4 (CHD)	Patient 5 (CHD)	Patient 6 (CHD)	Patient 7 (CHD)	Patient 8	Patient 9	Patient 10	Patient 11
D1S225	227 466	N/d	N/d	N/d	N/d	N/d	N/d	N/d	N/d	N/d	N/d	N/d
D1S251	228 021	del	N/d	N/I	N/d	N/d	N/d	N/I	N/d	N/d	N/d	N/d
D1S2709	228 337	N/I	N/d	N/d	N/d	N/I	N/d	N/I	N/I	N/d	N/d	N/d
D1S459	228 801	del	N/I	N/d	N/d	N/d	N/d	N/d	N/I	N/I	N/I	N/d
B804 (RP11-804P4)	229 460	del	N/I	N/I	N/d	N/I	N/d	N/d	N/d	N/d	N/d	N/d
B895 (RP5-895L5)	230 123	del	N/d	N/I	N/d	N/I	N/d	N/d	N/d	N/d	N/I	N/I
B781 (RP4-781K5)	231 047	del	N/d	N/d	N/I	N/d	N/d	N/d	N/d	N/I	N/d	N/d
D1S235	232 219	N/I	del	N/d	N/d	N/d	N/d	N/d	N/d	N/d	N/d	N/I
B940 (RP5-940F7)	232 415	N/I	del	N/d	N/d	N/d	N/I	N/d	N/d	N/I	N/d	N/d
D1S2850	232 956	N/I	N/I	N/I	N/d	N/I	N/d	N/d	N/d	N/d	-	N/d
B44 (RP11-44E24)	233 694	N/I	N/I	del	N/I	N/I	N/d	N/d	N/d	N/d	N/I	N/I
B541 (RP11-541M23)	233 983	del	del	N/I	N/d	N/I	N/d	N/d	N/I	N/d	N/d	N/d
B193 (RP11-193H5)	234 387	N/I	del	del	del	N/d	N/d	N/d	N/d	N/d	N/d	N/d
B307 (RP11-307O1)	235 533	N/I	del	del	del	N/d	N/d	N/d	N/d	N/d	N/d	N/d
D1S517	235 686	del	del	del	del	N/I	N/d	N/d	N/d	N/I	N/I	N/d
B359 (RP11-359A17)	235 850	del	del	del	del	N/d	N/d	N/d	N/d	N/d	N/d	N/d
D1S1149	236 473	N/d	del	del	del	N/I	N/d	N/d	N/I	N/d	N/d	del
B177 (RP11-177F11)	236 840	N/I	N/I	N/I	N/I	N/d	N/d	N/d	N/d	N/I	del	N/I
D1S2785	237 202	N/d	del	N/I	del	del	N/d	N/d	N/d	del	del	N/I
D1S180	237 691	N/d	N/I	N/I	N/I	N/I	del	N/d	N/d	del	del	del
D1S547	238 081	N/d	del	del	del	N/I	N/I	del	N/d	del	N/I	del
D1S2842	239 198	N/d	del	del	del	del	del	del	N/I	del	del	N/I
D1S1609	240 391	N/d	N/I	del	del	del	del	N/I	N/d	N/I	del	N/I
D1S2682	244 456	N/d	del	del	del	N/I	N/I	del	del	N/I	del	del
D1S3739	245 358	N/I	del	N/d	del	del	del	del	del	del	del	del

Table 3.4: A table summarising the results of microsatellite analysis carried out on patients 1-11. N/d - not deleted (heterozygous); del – deleted; N/I – not informative. Microsatellite analysis was carried out as before (section 3.3.1). Additional markers were obtained from the 1q44 region. Deleted markers and non-informative markers assumed to be deleted are highlighted in red. Patients 1-7 have a cardiac defect. Positional information was taken from http://www.ncbi.nlm.nih.gov/, as of 10/04.

3.3.3 Loss-of-heterozygosity analysis

A panel of 15 affected individuals with CHD with normal karyotypes, patients 12-28, were identified for LOH studies. All the patients had obstructive outflow tract defects and hypoplastic left heart. No parental DNA was available to identify non-inheritance of amplified alleles. Therefore only patients with a lymphoblastoid cell line available were chosen. This permitted metaphase chromosomes to be obtained for FISH analysis should a contiguous block of single allele amplified markers be identified. The microsatellite markers and the primers used were the same as for the previous chapter. The results of the loss-of heterozygosity analysis are shown in table 3.5. Three individuals, patients 13, 14 and 21, had a block of three or more contiguous polymorphic markers amplifying only a single allele. None of these were overlapping. Fluorescence *in situ* hybridisation was carried out to determine whether these were homozygous or hemizygous. BAC and PAC genomic clones mapping to the three block or markers were identified from the NCBI database. STS amplification confirmed that the correct clone had been received (data not shown). Patient 13 metaphase chromosomes were probed with RP4-804P4, patient 14 with RP11-477C23 and patient 21 with RP11-359A17. These were labelled with FITC by nick translation. A chromosome 1 plasmid marker was labelled in Texas Red. These were co-hybridised to patient metaphase chromosomes. FISH analysis showed that none of these patients were deleted.

	Genotype (bp)														
Locus	**Pt 12**	**Pt 13**	**Pt. 14**	**Pt. 15**	**Pt. 16**	**Pt. 17**	**Pt 18**	**Pt. 19**	**Pt. 20**	**Pt 21**	**Pt. 22**	**Pt. 23**	**Pt. 24**	**Pt. 25**	**Pt. 26**
D1S225	219/231	219/231	229/231	219/235	229/233	217/219	225/235	233/237	237	229/233	219/239	217/219	217/219	229	229
D1S251	261/275	261/263	257/259	259/263	257/267	257/267	253/257	259/267	255/267	259/267	267/269	257/267	259/267	259/267	253/261
D1S2709	193/195	193/195	191/195	195	195/197	195/197	199	193	195	195/197	195/197	193/197	195	197/199	195/197
D1S459	180/184	180/182	180	180/182	174/180	176/186	180	180/184	180/182	180	180	182	180/188	182/184	176/180
B804	1387/141	141	141	141	141	141/151	139/141	141	141	141/143	143	147	139/1411	137/141	141
B895	281/283	283	281/283	281/283	281/283	281/283	281/283	281/283	279/285	279/283	281/285	281/283	283/285	281/283	281/283
B781	240	230	230	232	232/244	240	230/244	230	240	232/242	240	240/242	228/232	228/240	242/244
D1S235	176	176/188	188/190	176/182	176	188/190	176/188	188	176/188	176/192	176/190	176/188	176/196	176/190	190/194
B940	321/323	319	311/323	319	319	319	319/321	315/319	317/319	317/319	319	319/323	321/323	307/319	317/323
D1S2850	145/147	145/147	147/149	149/151	151/155	149	147/149	149/151	147/151	139/147	147/151	147/151	149	151/153	151/153
B44	191	191/199	191/201	201	191	191/201	189/203	191	191	191/201	191/199	189/191	189/191	191	191
B541	287/289	293/297	287/293	283/287	285/287	293/295	287	281/291	281/291	279/287	281/293	287	291/295	283/285	287
B193	291/297	287/299	291	289/291	289/293	283/293	287/293	287/289	291/293	285/293	287/289	289	291/293	287/299	287/291
B307	320/336	318/336	322/336	324/340	316/332	320/326	338	320	322/330	322/342	308/334	320/336	314/334	320/330	308/340
D1S517	214/218	206/214	208/214	194/206	206/218	206	200/212	202/210	214/222	206/218	196/218	206/218	206/214	208/214	196/206
B359	223/235	223/235	223/231	225/233	227/233	231/235	227	227/235	231/233	233	235	233/239	227/231	223/227	223/227
D1S1149	377/385	383/389	387/391	383/385	379/393	371/385	385	383	383	387	379/385	383	377/385	377/379	379/383
B177	234	234	234/238	234/236	222/238	220/234	220/234	234/236	230/234	230	233/235	235/237	234	230/238	230/238
D1S2785	176/178	176/178	168/174	166/176	172/178	174/178	172/182	168/178	176/178	174	178/182	174/180	166/178	172/178	172/174
D1S180	169/173	169/173	163/173	165	169	167/187	161/163	169/171	173	171	173	163/173	165/169	163/187	163/187
D1S547	289	289	289/303	285/299	289/293	278/298	285/306	285/297	289	287/293	285/289	285/289	285/289	289	289
D1S2842	219/227	219/227	225/227	219/229	217/219	221	223/229	219	223/227	223/227	217/227	219/225	223/229	225/227	217/227
D1S1609	187	187	183/187	187/203	179/183	175/181	187	187/191	187	183/195	183/195	183/187	189/197	183/191	179/183
D1S423	163	163	163	163/165	165	163	163/165	165	163/165	-	163	165	165	163/165	163/165
D1S2215	247/255	247/255	255	255	251/255	253	255	252/259	247/255	-	247/255	255	249/257	255	251/253
D1S2682	117/137	117/137	117	119/135	117/135	117/133	119/133	117/137	133/137	133/137	137/139	117/137	119/137	119/133	117/137
D1S3739	186	186	186	186/190	186	156/186	180/186	186	186	186	156/190	186	186/190	186	182/186

Table 3.5: The results of loss-of-heterozygosity analysis carried out on patients 12-26. LOH analysis was carried out on a panel of 15 affected individuals. Patients 13, 14, and 21 had contiguous blocks of single allele amplifications (highlighted in blue).

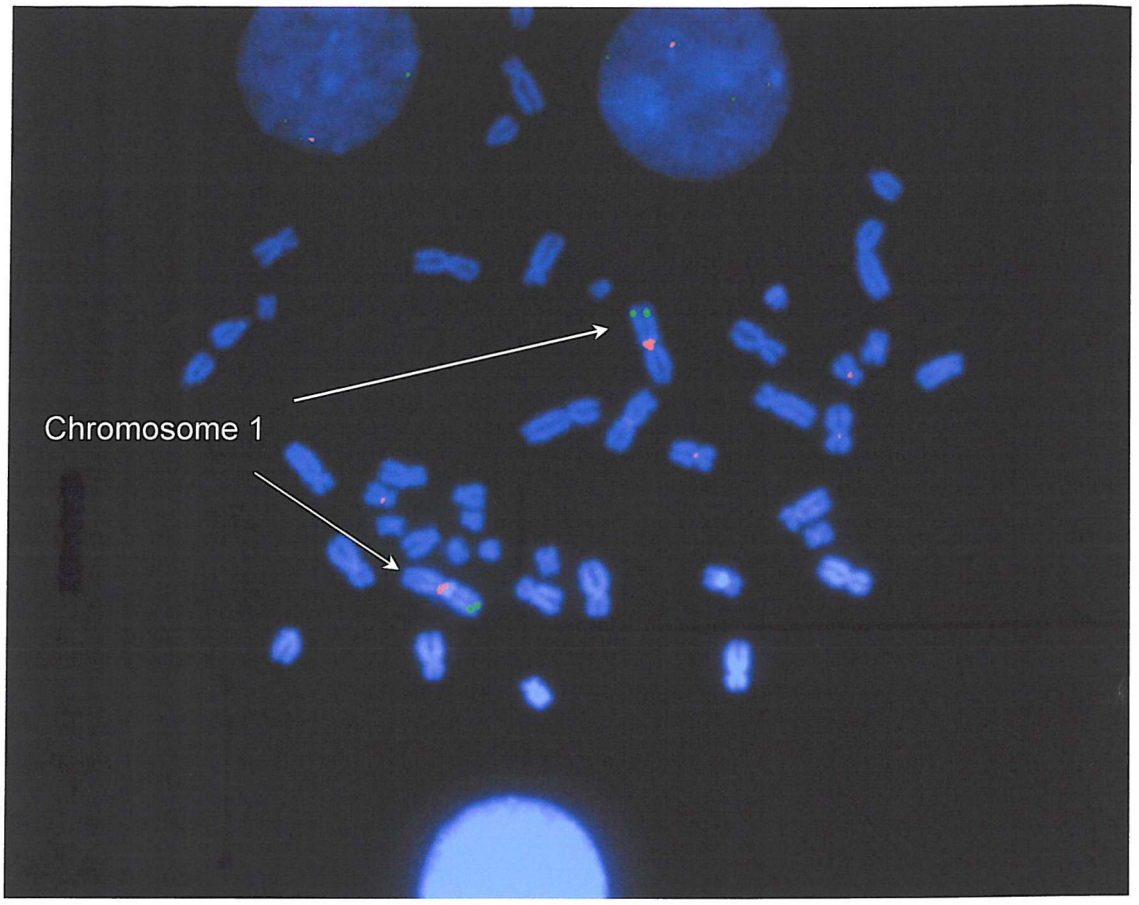

Fig 3.3: The FISH result of patient 13 metaphase chromosomes probed with PAC clone RP4-804P4. Patient metaphase chromosomes were stained with DAPI. A chromosome 1 plasmid marker labelled in Texas red hybridised to the chromosome 1 centromeres. Genomic clone RP4-804P4 labelled in FITC green hybridised to both chromosomes indicating that the patient was not deleted at this locus.

3.4 Discussion

Chromosomal imbalance within 1q42-44 can give rise to an array of cardiac defects, suggesting the presence of a crucial cardiac development gene within this region. No case reports to date have presented detailed molecular analysis of these patients. With a view to identifying candidate disease genes it was first necessary to establish a molecular critical CHD region. Microsatellite analysis was therefore carried out on eleven patients with known 1q42-44 deletions to establish a molecular CHD susceptibility interval. The data obtained from these studies may suggest that there are two CHD susceptibility loci within the 1q42-44 region. Analysis of an initial patient, patient 1, showed the presence of a 9Mb interstitial deletion. Subsequent investigations were carried out on a further 10 patients to reduce the critical region and therefore the list of candidate genes. Three CHD patients, 5-7, had distal deletions that did not overlap with that of patient 1. Patient 7 had the smallest deletion of approximately 7.5 Mb. Given that the deletion found in patient 7 was over a megabase from that identified in patient 1, this may indicate that there are two non-overlapping CHD critical regions, one located proximally defined by patient 1 and another distally defined by patient 7 (see fig 3.4). As suggested in the introduction (section 3.1), this may be supported by published case reports. Two case reports presented cytogenetic data on two CHD patients who were shown to have proximal interstitial deletions extending from 1q32-42 (Al-Awadi et al., 1986; Sarda et al., 1992). Others reported with more distal terminal deletions extending from 1q42.3/ 1q43 to the telomere (Garani et al., 1988; Johnson et al., 1985; Meinecke and Vogtel, 1987; Watson et al., 1986; Wright et al., 1986). De Vries *et al.* reported on two CHD patients, who were shown to have deletions within 1q44 (De Vries, Knight *et al.* 2001). However at this tine it is not possible to discount the possibility that a crucial cardiac developmental lies between these two critical regions, which is influenced by position effect when either region is deleted. Patients 2-4 had deletions that overlapped with that of patient 1. It was not however possible to reduce the proximal critical region, as all three had deletions extending into 1q44. The cardiac defects seen in these patients may therefore have been the result of haploinsufficiency of a distal critical gene. No specific cardiac phenotype was associated with either deletion region.

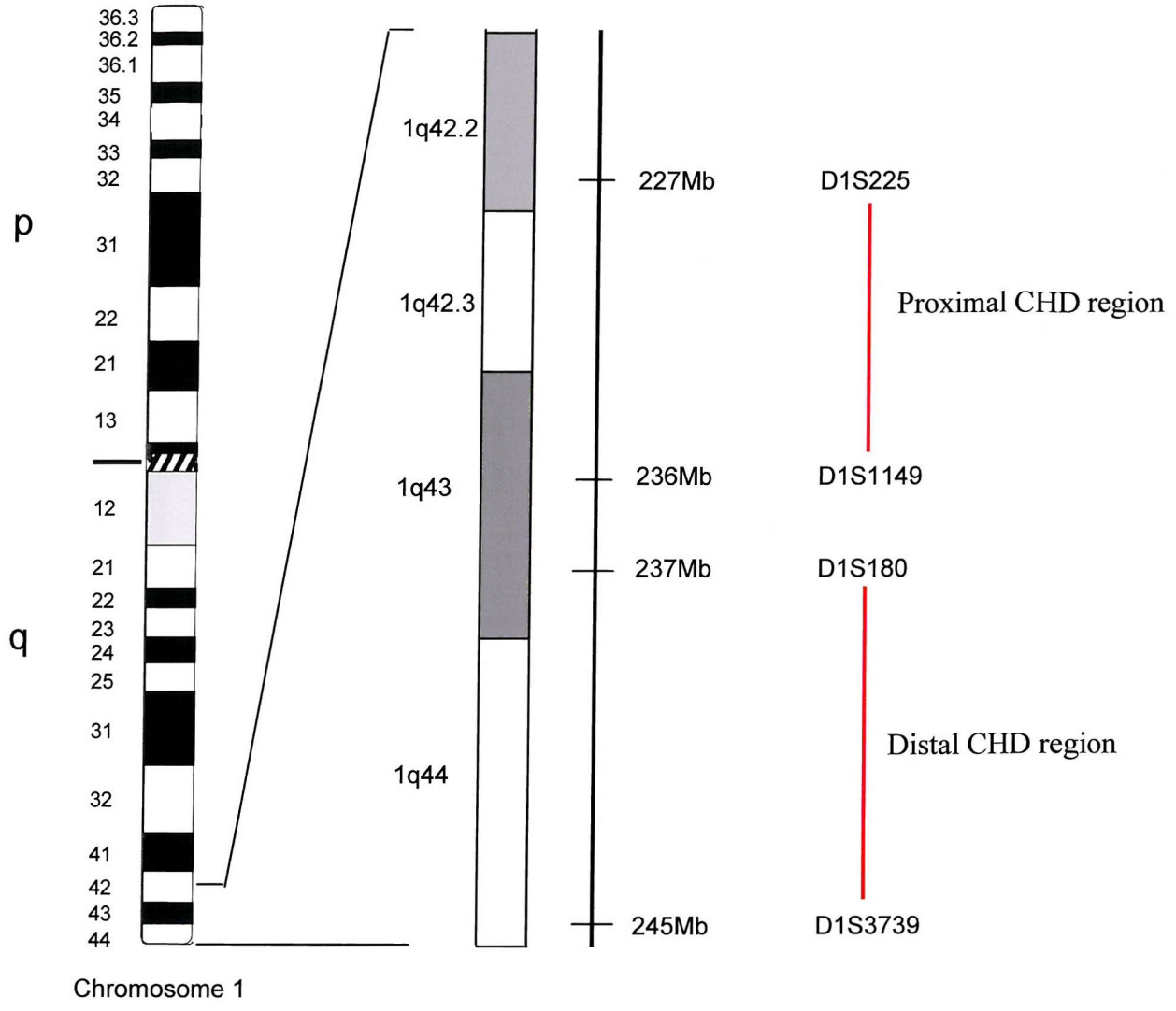

Fig 3.4: A schematic representation of the possible CHD susceptibility intervals within 1q42-44. Left of the diagram is an ideogram of chromosome 1 with inset to the right an enlarged view of the 1q42-44 region. To the right are the two proposed critical regions. The proximal CHD region, defined by patient 1, is bordered by DIS225 proximally and D1S1149 distally. The distal deletion region defined by patient 7 extends from D1S180 to the telomere. However it may be possible that a crucial cardiac developmental gene lies between these two regions which is affected by position affect when either critical region is deleted.

A panel of 15 patients with normal karyotypes were investigated for the presence of submicroscopic deletions within chromosome 1q42-44 which may have gone unidentified under routine diagnostic testing. This was undertaken using a combination of LOH analysis and FISH. Three patients with contiguous blocks of single allele markers were identified who were subsequently shown to be not deleted using FISH. Although microsatellite markers were chosen on the basis of heterozygosity scores for published markers, or on the presence of dinucleotide repeats for those designed in-house, a homozygous result arose with some frequency. In all three patients that underwent FISH investigations, the single alleles amplified as contiguous blocks were common to other patients. For example patient 13 was homozygous with an amplified allele of 141bp for marker B804. This allele size was common to thirteen of the fifteen patients, and seven of these were homozygous. Some of these markers were therefore not as polymorphic as expected. Without parental DNA, FISH confirmation was required to demonstrate the presence or absence of a deletion. The need for metaphase chromosomes for FISH analysis restricted the patient panel to those with an immortalised cell line established. In addition culturing of cell lines until sufficient cells are present for metaphase chromosomes to be obtained can take a long time. These limitations highlighted the need for a more efficient deletion detection methodology requiring no secondary confirmation using fluorescent *in situ* hybridisation, which would not only improve throughput but would also enable a greater number of patients to be investigated.

Chapter 4: Microarray comparative genomic hybridisation

4.1 Introduction

4.1.1 Development of microarray-comparative genomic hybridisation.

The development of comparative genomic hybridisation (CGH) was driven by the need for a technology that would allow genome-wide screening for alterations in DNA copy number, independent of the need for metaphase chromosomes from the test subject. Differentially labelled test DNA and control DNA sample are hybridised to a representation of the genome. Metaphase chromosomes were originally used as the hybridisation target. Repetitive sequence elements are blocked by the addition of competitor DNA, principally Cot-1 DNA. The resulting ratio of fluorescence intensities gives an indication of copy number of the test sample in comparison to that of the known normal DNA. In theory, a patient with a normal autosomal compliment number would give a ratio of 1, monosomy would be represented by a ratio of 0.5 and trisomy by 1.5. The principle limitation of early CGH studies was due to the use of metaphase chromosomes as a target substrate which limited the resolution to approximately 5 to 10 Mb. To overcome this, metaphase spreads were replaced by spotting arrays of large-insert genomic clones (BACs/PACs) as targets for hybridisation onto glass slides (Pinkel et al., 1998; Solinas-Toldo et al., 1997). The resolution of the array to detect chromosomal imbalances is therefore only limited by the size of the insert and the density of the arrayed clones. An additional advantage of array-CGH over metaphase-CGH is that the use of genomic clones with known map information allows rapid molecular definition.

Several studies have demonstrated the reliability and utility of array-CGH. In 1998 Pinkel *et al.* confirmed the sensitivity and quantitative capability of array-CGH for gene dosage measurements using an arrayed subset of genomic clones taken from chromosomes X and 22 (Pinkel et al., 1998). Single copy deletions and amplifications within tumour DNA were reliably detected to a resolution of approximately 40Kb. An equivalent level of genomic resolution was obtained in an investigation of neurofibromatosis type 2 (NF2; OMIM: 101 000; Bruder et al., 2001). A number of NF2 patients have constituitional deletions that remove part of or an entire copy of the NF2 gene. Bruder *et al* adopted an array-CGH approach to examine a 7Mb interval on chromosome 22 in the vicinity of the NF2 gene to determine the extent and frequency

of deletions within a panel of NF2 patients. A genomic array was designed around data obtained from previous molecular analysis of an NF2 deletion patient (Bruder et al., 2001; Bruder et al., 1999). 104 clones were obtained covering an estimated 90% of the 7.4 Mb interval. The arrays were validated using the DNA of the NF2 patient known to contain a 7Mb deletion within the area covered by the array. The normalised average ratio for the deleted clones was 0.59. The size of the deletion determined by microarray matched well with the previous analysis of the case (Bruder et al., 1999). Following validation, 116 NF2 patients were investigated for constitutional deletions. 24 were shown to have deletions ranging from 2.4 Mb to 40 Kb. Deletions could be detected to within a single cosmid. Array-CGH has also been successfully shown to have a genome-wide capability. In 2001, Snijders *et al.* developed a genome-wide DNA microarray that consisted of 2,460 BACs and P1 clones printed in triplicate distributed across the genome (Snijders et al., 2001). This provided an average resolution of approximately 1.4Mb across the genome. Subsequently Fiegler *et al.* produced arrays with 3,000 clones printed that were spaced at approximately 1Mb intervals across each chromosome arm (Fiegler et al., 2003a; Snijders et al., 2001). The use of such large numbers of genomic clones brings with it inherent difficulties in array construction.

4.1.2 Strategies for array construction

Initial reports of array CGH used whole BAC/PACs isolated from large bacterial cultures which were subsequently sonicated to fragment the DNA (Pinkel et al., 1998). Growing and processing large cultures is expensive and time-consuming, especially for genome-wide arrays, or those targeting large chromosomal regions, which require a high number of clones to ensure adequate genomic coverage. Additionally, BACs and PACs are single copy vectors, and DNA yield is therefore low. To overcome these difficulties, several strategies have been developed to amplify small amounts of cultured and purified clone DNA. Two of these, linker adapter-PCR and degenerate oligonucleotide primer-PCR (DOP-PCR) have given promising data (Fiegler et al., 2003a; Snijders et al., 2001).

Ligation-mediation PCR first involves digesting DNA with a frequent cutting restriction enzyme that results in fragments ranging from 200 – 2000bp. A universal

oligonucloetide adapter is then ligated to the fragment and this serves as the priming site for PCR amplification. This strategy was adopted by Snijders *et al.* in the construction of their genome wide array (Snijders et al., 2001). An alternative approach utilising an adapted DOP-PCR methodology has also been shown to give promising data. Conventional DOP-PCR completely amplifies a target DNA sample through the use of primers that can bind randomly (Telenius et al., 1992). However the use of the standard DOP-PCR primers compounds the problem of *Escherichia coli* contamination which is a common pollutant of DNA preparations from large-insert clones. *E. coli* contamination would increase non-specific binding to the arrayed clones, giving rise to increased background fluorescence and consequently unreliability in ratio estimation. This lead to the development of DOP-PCR primers that would preferentially amplify human over *E.coli* DNA (Fiegler et al., 2003a). Three 6mers were identified that arose with significantly more frequency in human than *E. coli* sequence and these were used to form the basis of the 3' sequence to three DOP-PCR primers (primer sequences can be found in appendix i). Subsequent testing confirmed their ability to preferentially amplify human over *E. coli* DNA, and showed that combining the products of all three primers provided the best target representation. A recent workshop conducted by the Wellcome Trust Sanger Institute sought to compare some of the strategies used for array construction (Carter et al., 2002). Fifteen laboratories participated in the workshop. Some groups spotted whole clone DNA. The majority spotted PCR-amplified clone using linker-adapter or DOP-PCR methodologies. Comparison of array performance clearly demonstrated that PCR-based methods were the most effective for array construction. No difference in performance between linker-adapter-PCR and DOP-PCR was highlighted.

Several factors may contribute to increased background fluorescence and a consequent decrease in array performance. This may arise if the cloned sequence under investigation shows significant homology to other regions throughout the genome, through the presence of low-copy repeats, *alu* repeats or pseudogenes. Prior to arraying, ideally, clones need to be FISH-mapped to metaphase chromosomes, to ensure, firstly that the correct clone has been received, and secondly that there is no cross-hybridisation to other chromosomes. Arraying of large genomic clones prevents pre-seletion of the target sequence. Recent studies have sought to PCR specific DNA sequence, to eliminate repeat sequences from the array (see section 6.2).

Bacteriophage contamination of genomic clones may also give rise to background fluorescence, requiring that each clone is checked to ensure that it is contamination free. Large quantities of Cot-1 DNA are used to block repetitive sequence. However many groups have commented on the variability between batches of Cot-1 in their ability to suppress hybridisation to repeat sequences (Carter et al., 2002)

4.1.3 Development of array-CGH for detecting deletions in patients with congenital malformation syndromes

Although array-CGH has principally been applied to the investigation of cancer genetics, it also has enormous potential in the study of congenital syndromes. High-resolution mapping of critical deletion regions is of importance for the subsequent discovery of disease-associated gene(s). Rauen *et al.* recently investigated an individual diagnosed with cardio-facio-cutaneous (CFC) syndrome, presenting further evidence for a CFC candidate region (Rauen et al., 2002). CFC patients present with characteristic craniofacial dysmorphlogy, cardiac defects and developmental delay. The cardiac defects vary, the most prevalent being pulmonary stenosis and atrial septal defects. Microarray analysis was carried out to obtain molecular data of a patient with a known 12q21.q22 deletion. Array-CGH analysis gave concordant data to that demonstrated by cytogenetic analysis. 3 BAC clones gave a deleted ratio, representing a deletion of between 14-19 Mb. Patients with a CFC phenotype, and a normal karyotype by standard cytogenetic banding are to be analysed in future attempts to reduce the disease region. A similar approach was adopted by Yu *et al.* to more rapidly identify breakpoints locations within 1p (Yu et al., 2003). Monosomy 1p36 is the most common terminal deletion syndrome (Heilstedt et al., 2003a). Heilstedt *et al.* reported the molecular characterisation of 60 patients with terminal deletions of 1p36, demonstrating that deletion sizes widely varied over a 10.5 Mb region within 1p36 (Heilstedt et al., 2003b). Yu *et al.* generated a BAC/PAC contig of distal 1p36 to develop a microarray to validate the application and accuracy of array CGH for detecting single-copy gains and losses (Yu et al., 2003). 97 clones were identified spanning the 10.5 Mb gap and printed on to glass slides. The arrays were used to interrogate 25 subjects with well characterised deletions of 1p36. A single copy-deletion was correctly identified in all 25 cell lines. In a single experiment array

CGH had correctly narrowed the breakpoint junctions of each subject to within approximately 100-300 kb.

Microarray technology has enormous potential within a positional candidate strategy as a high-throughput sensitive methodology for detecting submicroscopic deletions in order to refine critical disease intervals. A proof-of-principle preliminary series of array-CGH slides were therefore to be developed targetting the 8Mb proximal disease region defined by patient 1 identified in the previous chapter. These were to be validated with analysis of patient 1 DNA. BAC/PAC contigs covering the region were to be identified from NCBI (www.ncbi.nim.nih.gov). Clones were to be FISH-mapped to patient 1 metaphase chromosome to ensure that the correct clone had been received and that they were single-copy. Given that arrays can be designed to target numerous areas of interest throughout the genome, this was a technology that was not only of relevance to this thesis but to several ongoing projects within the laboratory, in particular the investigation into the 11q25 hypoplastic left heart region (Phillips et al., 2002). A tiling BAC/PAC contig of the 11q critical region had been previously identified and clones mapped using metaphase-FISH. Following extraction, DNA from the 1q and 11q regions were to be amplified using the modified DOP-PCR protocols and arraying methodologies designed at the Wellcome Trust Sanger Institute (Fiegler et al., 2003a). In addition clones covering the chromosome 1 with a resolution of 1 Mb and control *drosophila* clones were also arrayed (http://www.ensembl.org/Homo_sapiens/cytoview). These were obtained from collaborators at the Wellcome Trust Sanger Institute (see section 2.3.1.5).

4.2.Aims

1. To design and develop a preliminary series of microarray-CGH slides targeting the proximal deletion defined by patient 1.

2. To validate these slides using the DNA of patient 1.

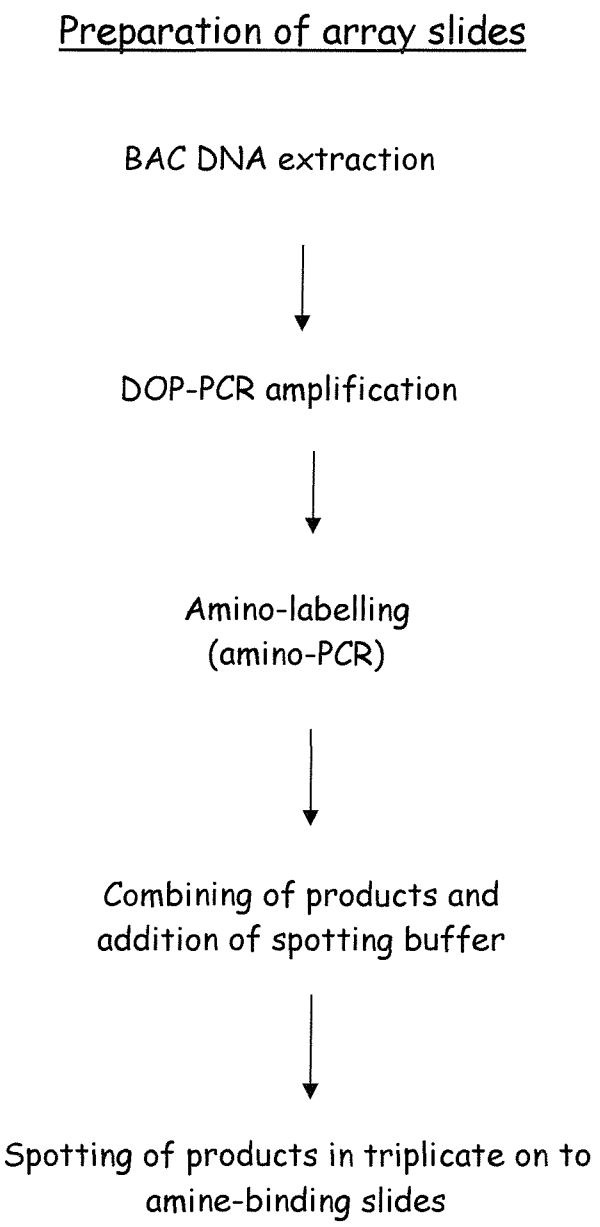

Fig 4.1: The steps of microarray slide preparation

(see section 2.4.1 for complete protocols)

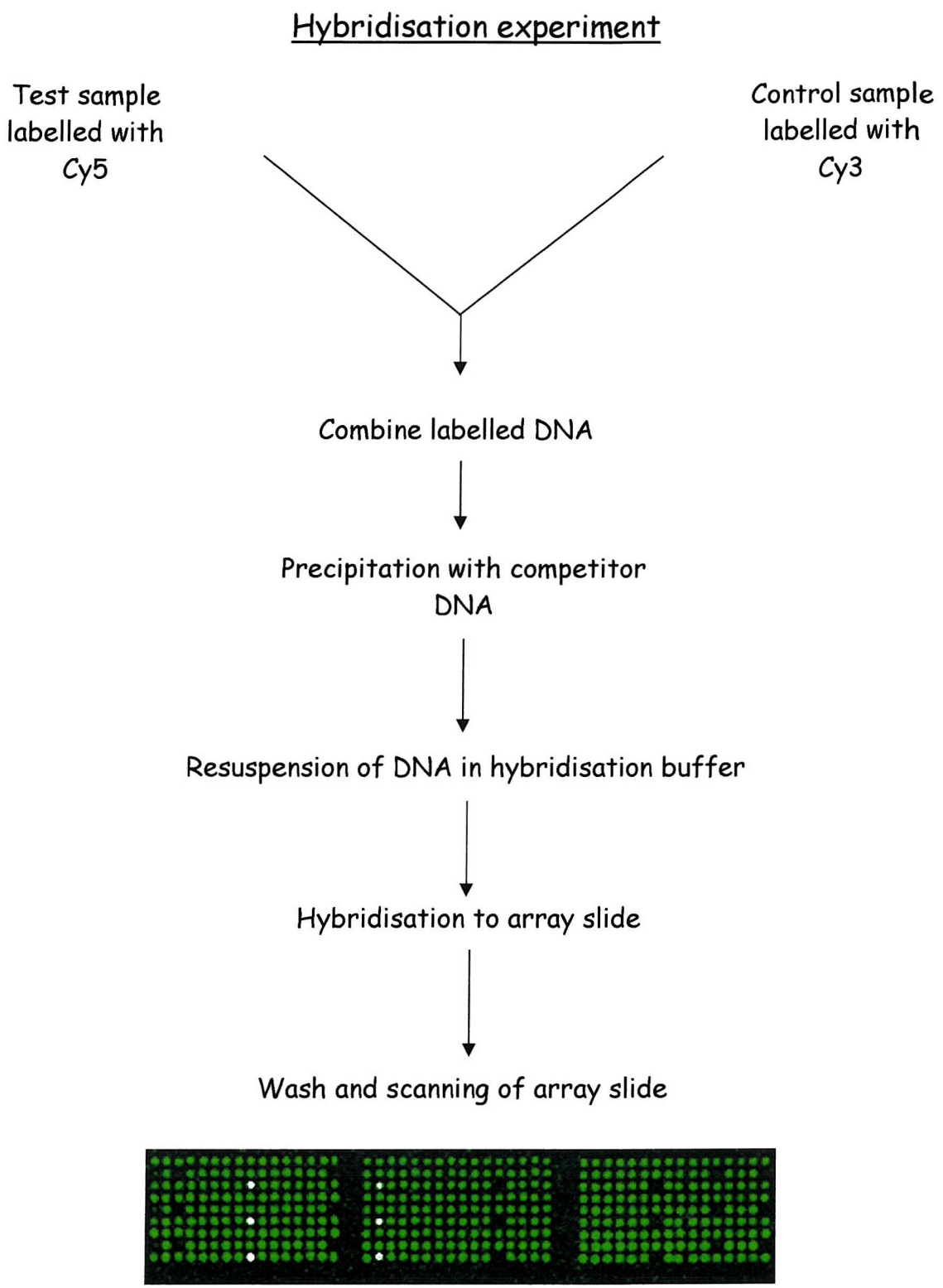

Fig 4.2: The steps of microarray-CGH hybridisation

(See section 2.4.2 for the full details of array hybridisation)

4.3 Results

52 clones were identified from the 1q42.3-43 8Mb critical region and obtained as agar stabs from the Wellcome Trust Sanger Institute. Each was tested for bacteriophage contamination using the methodology detailed in section 2.3.1.2. Contamination would result in lysis of the underlying bacterial lawn. None were shown to be contaminated. Clones were FISH-mapped to patient 1 metaphase chromosomes (see fig 4.3). BAC and PAC clones were labelled with green FITC by nick translation. Two clones mapped to different chromosomes and these were not used in the development of the array. FISH results have been summarised in table 4.2. 70 clones from the 11q HLH region had been previously identified and FISH-mapped. DNA extracted from BAC and PAC clones was amplified using the three modified DOP-PCR primers discussed in section 4.1.2. This was carried out in a sterile environment to eliminate the possibility of contamination which could result in background fluorescence during array hybridisation. Buffers were sterilised under UV-light. DOP-PCR was also carried out on negative controls containing water in place of DNA. Products were run on an agarose gel were and visualised on a UV-illuminometer to ensure the reactions had worked. No DNA was detected in the negative control samples, suggesting the products were contamination free. To facilitate attachment to the glass slides an amino group was added to the DOP-PCR products by PCR using a primer with an amino group attached to the 5' terminus. 40µl of the three amplified products for each clone were combined and array spotting buffer was added. The 1Mb chromosome 1 set and the control *Drosophila* clones were obtained prepared in array spotting buffer. Clones were arrayed in triplicate onto Motorola array slides at the Sanger Institute arraying facility.

The arrays were validated using the DNA of patient 1. Control DNA was labelled with Cy3 and patient 1 DNA in Cy5. The samples were cleaned in microspin columns to remove unincorporated fluorescent dye. The labelled DNA was then combined and precipitated with COT-1 DNA and herring sperm DNA to block repetitive sequence elements in the test and reference DNA. Following precipitation the DNA was re-suspended in hybridisation buffer. The DNA was applied to an array slide and left at 37°C for 48 hours. The slides was then washed to remove any unbound DNA and scanned

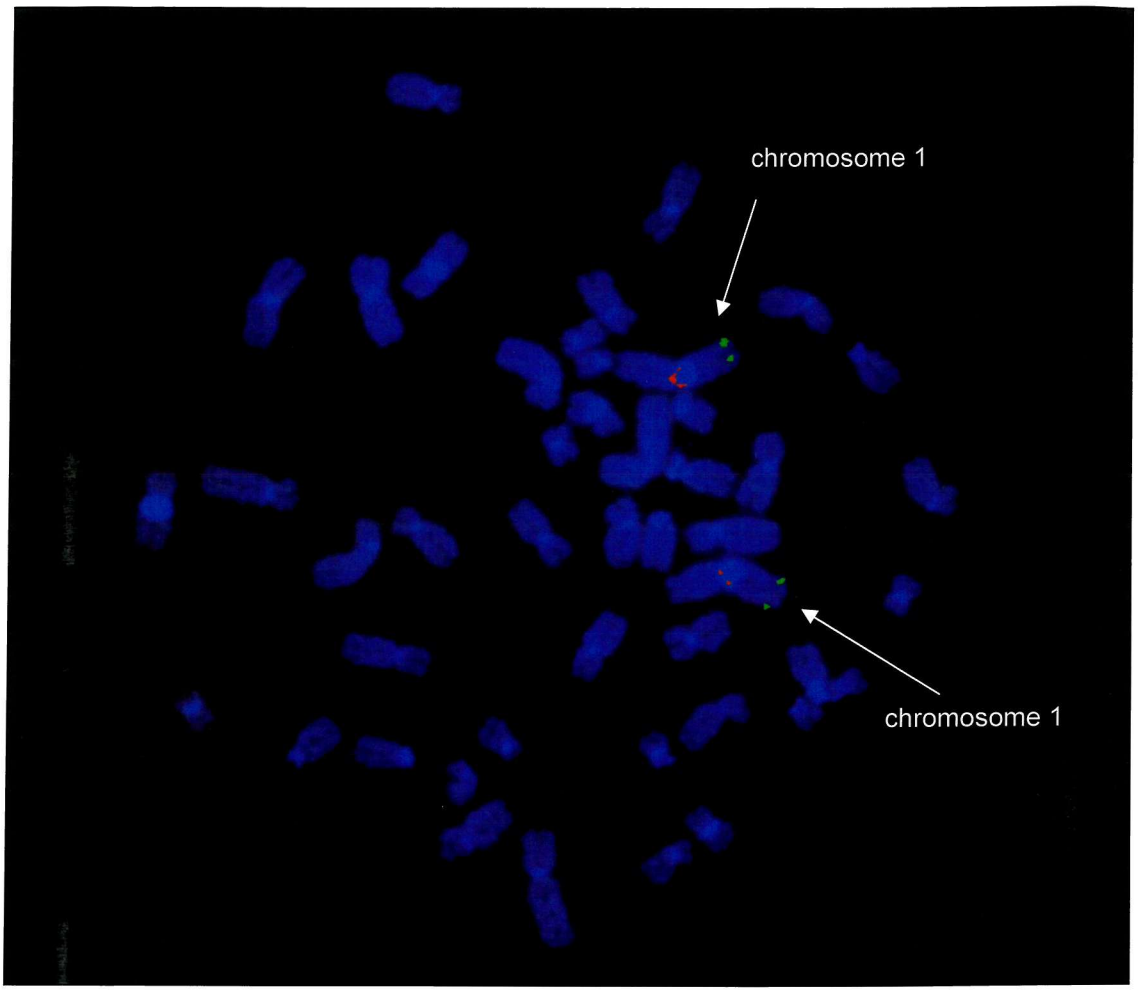

Fig 4.3: Patient 1 metaphase chromosomes probed with BAC clone RP11-876B10 to confirm localisation to chromosome 1q. Patient 1 metaphase chromosomes are stained with DAPI. A chromosome 1 plasmid marker, labelled with Texas Red was hybridised to the chromosome 1 centromeres. BAC clone, RP11-876B10, was labelled with FITC green and hybridised to the metaphase chromosomes to ensure that it mapped to the correct chromosomal region.

The Cy5 and Cy3 intensities for each spot were obtained following subtraction of local background fluorescence. Subsequent data analysis was carried out in Microsoft Excel. To correct for non-specific hybridisation to spotted DNA, the mean intensity of all the *Drosophila* clones was subtracted from both fluorochromes from all of the human clones. Spots with fluorescence intensities less than the mean *Drosophila* dye intensities were rejected from further analysis. The test/reference ratio for each spot were normalised by dividing each against the mean ratio for all the human cones. The normalised ratios were then accepted if they were within 20% of the median ratio for the triplicate. If one of the three was outside of this criteria, it was discarded from further analysis. If two were outside, all three were discarded. The mean ratio was then calculated for each of the accepted triplicate ratios.

Data was obtained which demonstrated that these slides could detect the 8Mb deletion found in patient 1. The results are summarised in table 4.1 and are displayed graphically in fig 4.4. The average ratio for the deleted clones was 0.71 (SD 0.012). The majority of the non-deleted clones gave a ratio of above 1. Six clones obtained from the Sanger Institute mapped to the deleted region. Four of the six gave a ratio of 0.6. Of the other two, clone, BA528D17 gave a ratio of 1.143 and no reliable data was obtained from clone DJ940F7. Clone RP11-87P4 was included as part of the 1q42-43 contig as well as in the chromosome 1 Sanger set. The fluorescent ratio for both was 0.621 and 0.613 respectively, which suggests reliability between target hybridisations. Seven clones from the target region, did not give reproducible data between the triplicate clones and as a consequence these were discarded from further analysis. The deletion data obtained from the microarray matches that obtained from the microsatellite analysis (section 3.2) and the FISH analysis of patient 1. The results for those mapping outside the deleted region can be seen in fig 4.4.

Clone identity	NCBI position (kb)	FISH Result	Array ratio (Cy5:Cy3)	SD
RP11-99J16	227 197	not deleted	1.093	0.024
RP5-876B10	227 713	not deleted	n/a	n/a
RP11-295G20	227 851	not deleted	n/a	n/a
RP11-17H4	228 003	deleted	0.788	0.027
RP11-98O1	228 070	deleted	0.766	0.015
RP11-399B22	228 223	deleted	0.563	0.007
RP5-823D19	228 882	deleted	0.558	0.023
RP11-87P4	229 024	deleted	0.621	0.016
BA87P4	229 024	-	0.613	0.000
RP11-286F20	229 207	deleted	n/a	n/a
RP4-659I19	229 305	deleted	n/a	n/a
RP4-804P4	229 460	deleted	0.611	0.009
RP5-893M4	229 561	deleted	0.649	0.001
RP5-1016N21	229 562	deleted	0.59	0.008
RP11-109B17	229 712	deleted	0.622	0.009
RP5-862P8	229 729	deleted	0.644	0.002
RP11-285B4	229 846	deleted	0.713	0.009
BA528D17	229 940	-	1.143	0.041
RP11-274P19	230 001	deleted	0.747	0.018
RP4-550F15	230 052	deleted	0.776	0.004
RP11-301G21	230 230	deleted	0.695	0.003
RP5-855F14	230 963	deleted	0.644	0.011
RP11-443B7	231 290	deleted	0.739	0.016
RP11-382D8	231 707	deleted	0.72	0.021
RP11-517I24	231 847	deleted	0.784	0.010
RP11-293G6	231 896	deleted	0.867	0.013
RP11-365D9	232 055	deleted	0.75	0.011
BA509I1	232 131	-	0.666	0.004
RP5-1043F6	232 209	deleted	0.638	0.000
DJ940F7	232 415	-	n/a	n/a
RP11-182B22	233 330	deleted	0.736	0.006
RP11-192B19	233 501	deleted	0.775	0.022
RP11-47A4	233 529	deleted	0.668	0.010
RP11-44E24	233 694	deleted	0.693	0.020
BA214M7	233 801	-	0.666	0.004
RP11-302J17	233 985	deleted	0.569	0.007
RP11-148L4	234 445	deleted	0.853	0.010
RP11-38P16	234 627	deleted	0.712	0.008
RP11-236L13	234 671	deleted	0.673	0.014
RP11-136B18	234 749	deleted	0.658	0.020
RP11-433N10	234 898	deleted	n/a	n/a
BA433N10	234 898	-	0.699	0.016
RP11-177F15	235 021	deleted	0.696	0.020
RP11-362H12	235 159	deleted	n/a	n/a
RP11-63B15	235 392	deleted	0.78	0.020
RP11-498G23	235 675	not deleted	1.012	0.007
RP11-96F14	236 715	not deleted	1.04	0.004
RP11-177F11	236 840	not deleted	1.023	0.012
RP11-467120	236 965	not deleted	1.057	0.026

Table 4.1: Table legend located on the following page

Table 4.1 (overleaf): A table showing the results of FISH and array-CGH analysis of patient 1. Genomic clones prefixed with BA- or DJ- formed part of the chromosome 1 Sanger set. There is no FISH data available on these. The position of clones is shown relative to the NCBI map (genomic sequence as of 10/04). The deletion results obtained from the array analysis match those obtained from the metaphase FISH. No data was obtained from 7 of the clones (n/a), as at least two of the triplicate clones were outside the accepted median ratio. Clone BA528D17 gave a ratio of 1.143, an unexpected result as it maps to the deleted region. This result may be due to repeat polymorphisms within the clone which are homologous to other genomic regions, or possibly due to the wrong clone being printed.

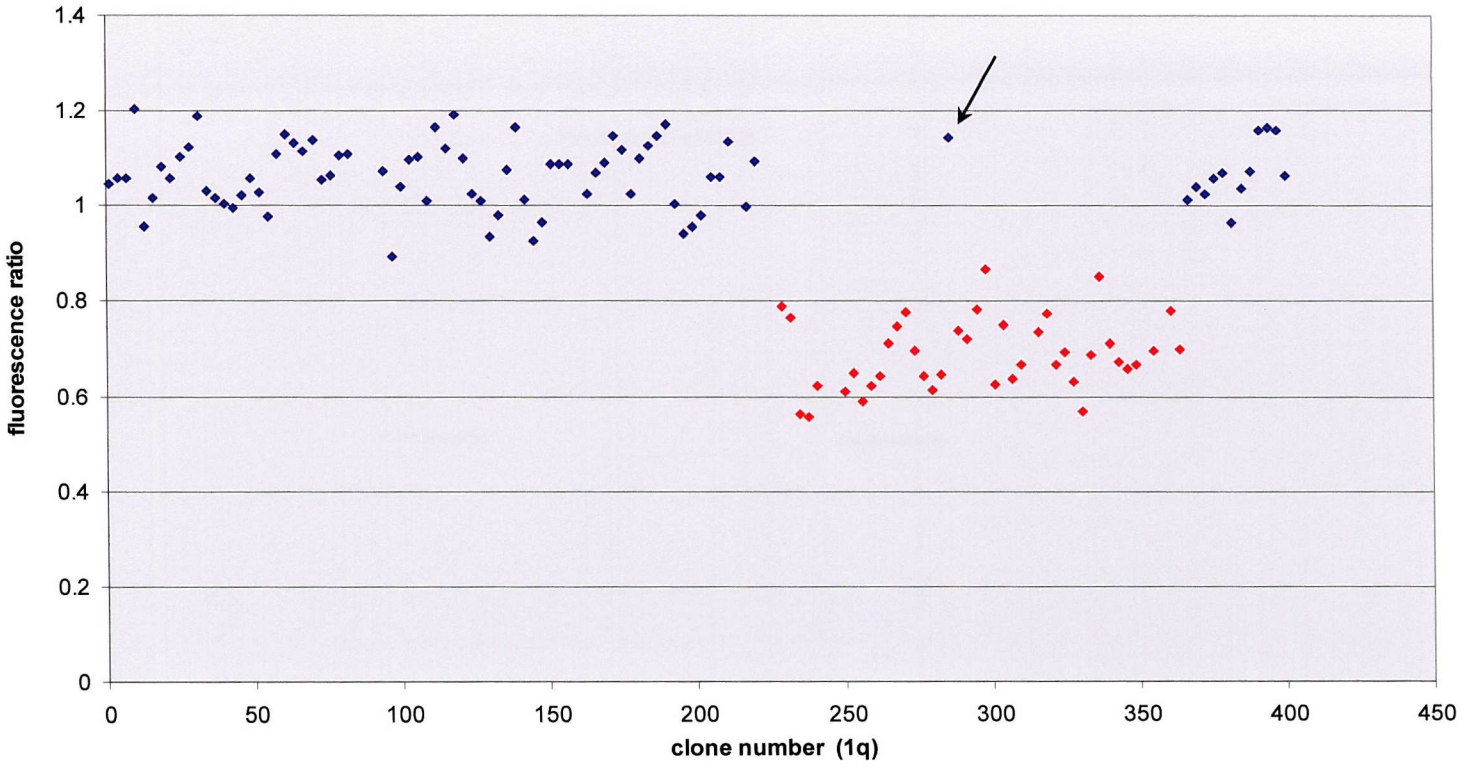

Fig 4.4: Array-CGH results of patient 1. Copy number variation of genomic clones from the chromosome 1 centromere to the 1q terminus (left to right of graph), showing the presence of a heterozygous interstitial deletion of patient1. The mean fluorescence ratio of the triplicate spots is plotted relative to the known order of the genomic clones along the length of 1q. Deleted clones are highlighted in red. This data is consistent with that obtained from microsatellite and FISH analysis. Clone BA528D17 is indicated by the arrow.

4.4 Discussion

Numerous studies have demonstrated the reliability and sensitivity of microarray-CGH in detecting chromosomal deletions to a resolution of 40kb. Given this, and the limitations identified in the previous chapter in searching for submicroscopic deletions by LOH analysis, an array-CGH approach would provide a higher throughput and more sensitive means of screening a patient panel for 1q deletions. In addition it would remove the need for metaphase chromosomes from the test subject, increasing the number of patients that could be analysed.

A preliminary series of slides were therefore to be developed targeting the proximal critical deletion region and validated through analysis of a previously characterised patient. 52 clones were obtained from the proximal critical CHD region. At the time of array design the sequence for chromosome 1 was incomplete and as a consequence complete coverage of this region was not possible. In addition map order and position changed with some regularity. Subsequent to printing, five of the genomic clones have since been abandoned from the NCBI database and positional information is no longer available on these. With improved sequence now available, genomic clones can be chosen with greater confidence that the position and order will not change. Two of the 52 clones were subsequently discarded from further use as they were shown to localise to different chromosomes, which confirmed the need to FISH-map clones before arraying. As minprep methodologies yield only small quantities of DNA, DOP-PCR was carried out using specifically modified primers designed to preferentially amplify human over *E. Coli* DNA. DOP-PCR was an efficient means of amplifying DNA which produced DNA product that subsequently went on to give reliable data.

The arrays were validated through analysis of DNA from patient 1. The ratios obtained were higher than that expected. Normal chromosomal compliment would give a ratio of 1, monosomy would give 0.5. The average for the deleted clones was 0.71. Others have also reported ratios that differed from the expected (Bruder et al., 2001). It may be the result of labelling bias with Cy5 being able to label more efficiently, which may be due to differences in the initial quality of the DNA samples to be labelled. Alternatively it may be due to incomplete suppression of repetitive

sequences with Cot-1 DNA or possibly autofluoresence of the target clones. No data was obtained from seven clones within the critical region. This was because at least two of the triplicate gave ratios that were outside the accepted. The data was therefore discarded. The variation of intensity within clone triplicates may be the due to the presence of repeat polymorphisms. Clone, BA528D17 gave an unexpected ratio. This clone maps to the deleted region and following hybridisation gave result of 1.143. Subsequent FISH analysis using this clone carried out by the Developmental Genetics Group at the University of Southampton confirmed that the patient was deleted at this locus. This therefore may suggest that either the wrong clone has been printed or that it has significant homology to numerous regions throughout the genome. The array result of patient 1 matched that obtained from FISH analysis, which suggests that the data is reliable. The clone DNA prepared in this investigation is still available to enable further slides to be printed in the future. Further experiments will determine the hybridisation characteristics of these clones and demonstrate whether these results occur consistently.

Chapter 5: Investigation of Nidogen-1

5.1 Introduction

5.1.1 Identifying the candidate genes within the critical region

Genes mapping to 1q42-ter were identified using the NCBI database (http://www.ncbi.nlm.nih.gov). A full list of genes and ESTs from this interval can be found in appendix ii. Genes were assessed for candidacy using the criteria laid out in section 1.5.3. A strong candidate gene, nidogen-1 was identified in the proximal critical deletion region. The majority of genes within the terminal 1q44 region were at the time of investigation uncharacterised with no known function. No obvious candidate CHD gene mapped to the terminal 1q44 region.

5.1.2 Structure and function of nidogen-1

Nidogen-1, alternatively called entactin, is located at 232 466 152 – 232 554 503kb within chromosome 1q43. The gene consists of 20 exons spanning more than 90kb. Exon 20 comprises a large 3' UTR. Full sequence and exon information can be found in appendix iii. Nidogen-1 is a 150-kDa glycoprotein, consisting of 1247 amino acids. The molecule consists of three globular domains G1-G3. G1 and G2 are separated by a flexible, protease-sensitive linker. G2 and G3 are connected via a rigid tandem of epidermal growth factors (EGF)-like domains (Fox et al., 1991). Sequence comparison of murine and human cDNAs revealed 84% and 85 % similarity at the nucleotide and the amino acid level respectively (Zimmermann et al., 1995).

Nidogen-1 is a component of basement membranes (BM). These structures are highly specialised elements of the extracellular matrix. Located at the interface between epithelia and mesenchyme in most tissues, these membranes serve as selective barriers and as scaffolds for adherent cells. The individual components of the BM can regulate cell growth, differentiation and migration, and that they can influence tissue development and repair. Made up of numerous proteins and proteoglycans, composition of the basement membrane varies depending on tissue type. All basement membranes have four major constituents; collagen type IV, laminin-1, perlecan and nidogen-1. Nidogen-1 has a high *in vitro* binding activity for laminin-1 (fox 1991), and to collagen type IV (Aumailley et al., 1989; Fox et al., 1991).

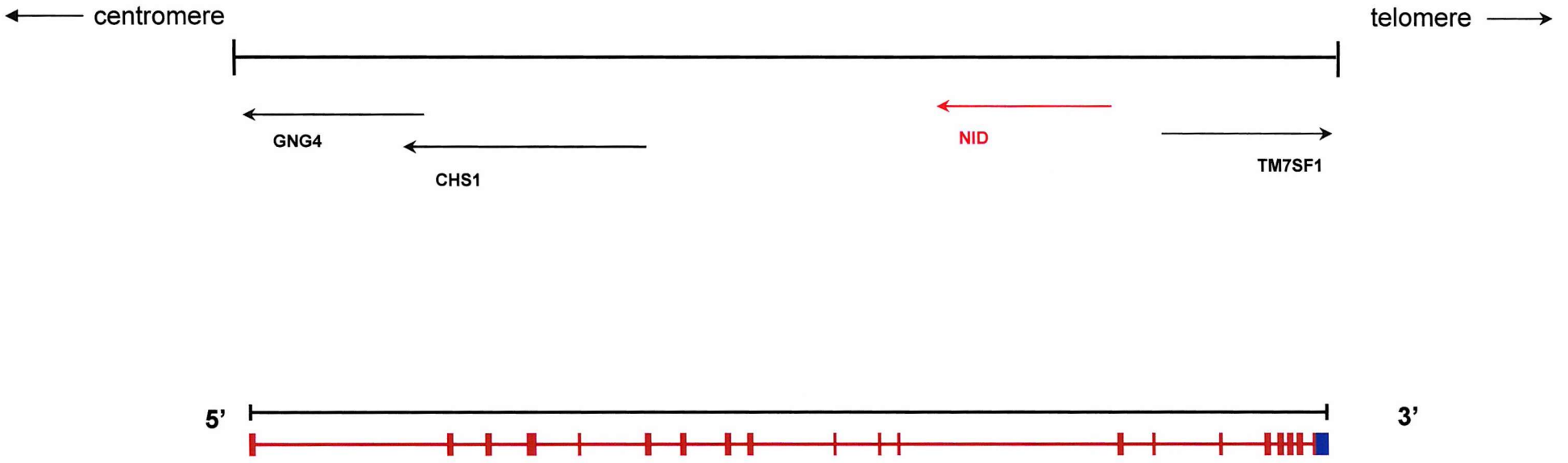

Fig 5.1: The genomic context and structure of nidogen-1. The red boxes represent nidogen-1 exons, 1 -20. The blue box represents the 3' untranslated region within exon 20. Adapted from http://www.ncbi.nlm.nih.gov/.

It has been shown to mediate the formation of structural complexes between laminin-1 and collagen type IV *in vitro* which led to the suggestion that it serves as a crucial link stabilizing the laminin-1 and collagen type IV networks within the membrane (Aumailley et al., 1993; Timpl and Brown, 1996). Antibodies that inhibit the laminin-1-nidogen-1 interaction perturbed epithelial branching morphogenesis in organ culture of lung, kidney and salivary glands (Ekblom et al., 1994). Perlecan binds with high affinity to the G2 domain of nidogen-1 (Hopf et al., 2001). Minor components of the basement membrane are present in a tissue specific fashion.

5.1.3 Candidacy of nidogen-1 for a role in CHD

5.1.3.1 The basement membrane and cardiac development

Basement membranes may play an important role in cardiac development. Two basement membranes, a larger myocardially derived basement membrane and an endothelially derived basement membrane, are found within the cardiac jelly (Kitten et al., 1987). These are thought to play a crucial role in epithelial-to-mesenchymal transformation at the sites of endocardial cushion development (Mjaatvedt et al., 1991). At the onset of EMT, fibronectin is deposited in the myocardial basement membrane of the cardiac jelly within the outflow tract and atrioventricular canal regions (Kitten et al., 1987).

Knock-out studies have shown that components of the BM are crucial for cardiac development. Deletion of laminin in *Drosophila* resulted in incorrect morphogenesis of the dorsal vessel, the equivalent of the vertebrate heart (Yarnitzky and Volk, 1995). Perlecan-null embryos gave rise to transposition of the great arteries (Costell et al., 2002). Interestingly, the human orthologs of these genes map to regions deletions within which can cause CHD. A recent study in mice found that collagen IV was essential for basement membrane stability (Poschl et al., 2004). The major basement membrane constituents, laminin, nidogen-1 and perlecan are deposited correctly in collagen IV null embryos, although weaker expression staining for these components was observed in some instances. Lethality in all collagen IV deficient embryos occurred at a later stage of development. This suggested that collagen IV was necessary for maintenance of the integrity and function of the basement membranes, but was dispensible for BM formation. It was postulated that collagen IV has an

essential role to maintain the structural integrity of the membrane under increasing physical demands. Vascular bleeding in the heart and aorta was seen in these mice and may be the result of insufficient cell-cell contacts arising from basement membrane disintegration. Correct BM development is therefore essential especially within organs of high mechanical stress, such as the heart

Administration of retinoic acid has been shown to cause transposition of the great arteries in mice (Yasui et al., 1995). The endocardial ridges in the outflow tract of these animals are hypoplastic. Alterations in the components of the basement membranes were found in cardiac tissue of retinoic acid treated mice and it was suggested that the changes in BM composition may be responsible for the hypoplastic ridges seen in these animals (Nakajima et al., 1997).

5.1.3.2 Nidogen-1 interaction with the elastic fibres

Nidogen-1 may interact indirectly with components of the elastic fibres within the outflow tract. The elastic fibres within the walls of the arteries are composed of two principal components; an amorphous elastin core consisting mainly of elastin, and a microfibril component consisting of fibrillin-1. The elastic fibres give the vessel walls the necessary elasticity to stretch and recoil following ventricular contraction. Mutations in fibrillin-1 and elastin can cause Marfan's Syndrome and SVAS respectively (see section 1.3). This interaction between nidogen-1 and the elastic fibres may take place through members of the fibulin family. There are now five known members of the fibulin family. The fibulins share tandem arrays of calcium-binding consensus sequences and have a diverse repertoire of interaction potentials, which makes them widespread components of the ECM (reviewed in Timpl et al., 2003). Some have been shown to play a crucial role in embryonic development. Fibulin-5 knock-out mice showed defects in heart, aorta and lung development (Nakamura et al., 2002). Histological analysis indicated defective development of the elastic fibre. Studies have demonstrated a strong affinity for of fibulins 1 and 2 with nidogen-1 (Sasaki et al., 1995b). Fibulin-1 is widely expressed in the ECM of various organs, where it associates with matrix fibres or with basement membranes. Fibulin-2 shows a more restricted expression pattern. Early in mouse development fibulin-2 expression suddenly increases in the endocardial cells, and in the transformed mesenchymal cells in the extracellular matrix of the cardiac outflow tract and the

endocardial cushions (Tsuda et al., 2001; Zhang et al., 1995). Later in development *in situ* hybridisation revealed fibulin-2 transcript specifically expressed in the mesenchymal cells of the semi-lunar and atrioventricular valves, the walls of the great vessels, the interatrial septum and the epicardium. Fibulin-1 and -2 have also been shown to interact with components of the elastic fibre. Both fibulin-1 and -2 were shown to bind the elastin precursor, tropeoelastin, although the affinity between fibulin-2 and tropeoelastin was higher (Sasaki et al., 1999). Fibulin-2 was also shown to interact with fibrillin-1 (Reinhardt et al., 1996). It therefore may be possible that fibulin-2 may link the elastin and fibrillin components of the elastic fibre to the basement membrane through its interaction with nidogen-1.

5.1.3.3 Nidogen-1 and the endocardial cushions

Nidogen-1 may play a role in the development of the cardiac valves and septa. As discussed previously the basement membrane is thought to play an important role in the deposition of fibronectin within the developing endocardial cushions during epithelial to mesenchymal transformation. The proteoglycan versican is expressed throughout development in the endocardial cushions and valve (Henderson and Copp, 1998; Olin et al., 2001). Knock-out studies of versican in mice displayed an absence of the endocardial cushion swellings leading to the conclusion that versican is required for cushion development (Mjaatvedt et al., 1998). Studies have shown that versican binds directly with fibulin-2 (Olin et al., 2001). It is therefore possible that nidogen-1 could link components of the endocardial cushions to the basement membrane though intermediaries such as fibulin-2.

5.1.3.4 Nidogen-1 mouse knock-out

Given the evidence indicating that nidogen-1 has a fundamental role in basement membrane biology and in organogenesis, the generation of nidogen-1 deficient mice with no overt anatomical pathology was perhaps a surprising result (Murshed et al., 2000). No ultrastructural changes to the basement membranes were observed. Tissue *in situ* expression staining for nidogen-1 found no signal in homozygous null mice. However increased nidogen-2 expression in cardiac tissue of nidogen-1 knock-out mice was observed and was reminiscent of nidogen-1 expression patterns observed in wild-type. This lead to a suggestion of genetic redundancy with nidogen-2 being upregulated to perform the functions of nidogen-1 in its absence. However crucial

differences in the binding affinities of nidogen-1 and nidogen-2 in human suggest that whilst nidogen-2 may be able to compensate for nidogen-1 in mice, this may not be the case in a human setting. Mouse and human nidogen-1 were previously shown to bind strongly to the laminin-1 γ1 chain and fibulin-2 (Mayer et al., 1993; Sasaki et al., 1995a). Human nidogen-2 however showed only a strong binding affinity for collagen iv and perlecan, and was a 100- to 1000-fold weaker ligand for laminin-1 (Kohfeldt et al., 1998). In addition, no significant binding of human nidogen-2 was observed to fibulin-1 or fibulin-2. This is in contrast to a distinct binding of both fibulins to nidogen-1. The binding repertoire of mouse nidogen-2 was recently investigated in comparison to nidogen-1 (Salmivirta et al., 2002). Both nidogens bound equally well to collagen IV and to fibulin-2. In addition nidogen-2 showed an only 8-fold lower activity compared to nidogen-1 in binding laminin-1. These findings may suggest that in the mouse system nidogen-2 may be able to compensate for absence of nidogen-1 giving rise to a mouse with no abnormalities. In humans, whilst nidogen-2 may be up-regulated in the absence of nidogen-1, it may not be able to bind crucially to laminin-1 and fibulin-2.

5.1.4 Investigating nidogen-1 for a role in congenital heart disease

Mutations in the nidogen-1 gene could result in disruption to interactions between the basement membrane and components of the extracellular matrix, in particular fibulin-2. Similar to elastin mutations and SVAS, aberrant development of the elastic fibre could be a result, giving rise to the outflow tract stenotic lesions, which is seen in cases of 1q monosomy. It also may be possible that nidogen-1 mutations could affect the fibulin-2 - versican interactions within the endocardial cushions resulting in the septal abnormalities consistently seen in these patients. Nidogen-1 is therefore a legitimate target to investigate for a potential role in congenital heart disease.

It was first necessary to confirm that nidogen-1 was expressed in cardiac tissue during human fetal development. The most straightforward method for achieving this was reverse transcriptase-PCR. Once expression had been confirmed by RT-PCR, immunohistochemistry analysis could be carried out to give a more precise indication of where nidogen-1 was expressed in the human fetal heart, and ascertain whether the expression pattern fits with the cardiac lesions seen in 1q deletion patients.

Immunolocalisation studies could also be attempted for laminin-1 and collagen type IV. To show that nidogen-1 does have a role in CHD aetiology, it is necessary to demonstrate the presence of pathogenic mutations. A panel of patients was therefore investigated to identify sequence variations. This was carried out by denaturing High Performance Liquid Chromatography (dHPLC) and direct sequencing. dHPLC is a highly sensitive, rapid means of screening for DNA sequence variations by the differential retention of homo- and hetero-duplex DNA. Sequence mismatches are represented by dual peaks on the dHPLC trace. The nature of any mismatches identified was to be determined by sequencing analysis using an automated ABI 377 automated sequencer.

5.2 Aims

1. To carry out expression analysis of nidogen-1 within human fetal heart using RT-PCR and immunhistochemisty analysis. Immunolocalisation studies were also to be carried out for laminin-1 and collagen type IV on human fetal sections.

2. To screen a panel of CHD patients for mutations within nidogen-1 using a combination of dHPLC analysis and direct sequencing.

5.3 Results

5.3.1 Expression analysis of nidogen-1

5.3.1.1 RT-PCR analysis

In order to confirm that nidogen-1 was expressed during cardiogenesis, RT-PCR was performed on mRNA from an 8 week human fetal heart. Primers were designed that were specific to nidogen-1 and these spanned an exon boundary. The forward primer was located in exon 3 and the reverse in exon 4 giving a product size of 212bp. A control experiment was carried out in parallel using primers to amplify the house-keeping gene, GAPDH. This would yield a product of 959bp and these primers were not intron spanning. mRNA was extracted from human fetal heart, liver, lung, pancreas and placenta. Any residual genomic DNA was removed following treatment with DNase. First strand cDNA synthesis was carried on the mRNA samples as detailed in section 2.4.3. Strong nidogen-1 expression was detected in all cDNA samples tested following PCR (see fig 5.1). The strongest expression was found in fetal heart. No amplification was seen in control samples of genomic DNA and pure water.

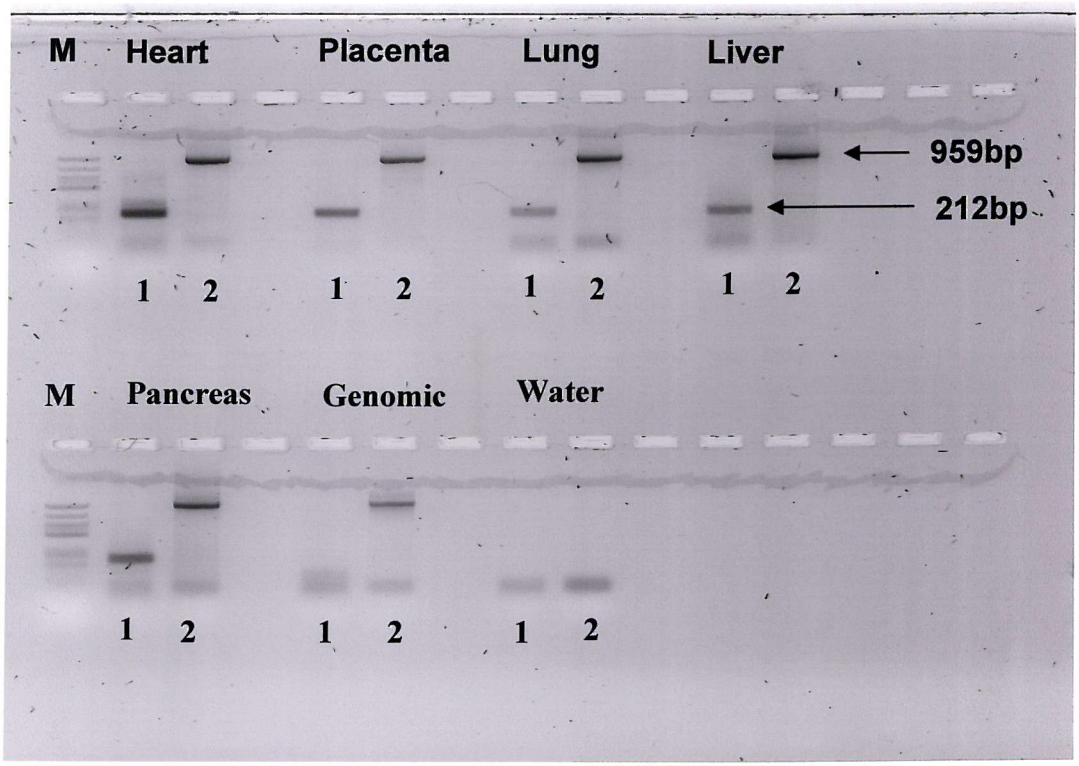

Fig 5.2 Results of PCR carried out on cDNA obtained from fetal tissue samples using primers to detect nidogen-1 expression. A band of 212bp represents nidogen-1 product. Nidogen-1 expression was detected in all fetal tissue samples; the strongest expression was seen in fetal heart. No expression was detected in genomic DNA as the primers span intron 3 which is over 3kb in length (see appendix iii). A control set of primers which amplify the house-keeping gene, GAPDH, were also used. GAPDH product is 959 bp in length. GAPDH primers had been designed within a single exon, which explains why the same size product is seen in the genomic and the cDNA samples. M-marker; 1 - nidogen-1; 2 – GAPDH

5.3.1.2 Immunohistochemisty analysis

Immunohistochemisty analysis was carried on an eight week human fetal heart sections using antibodies to nidogen-1, laminin-1 and collagen type IV. Dissection of the tissue sections and embedding into paraffin was carried out by the Developmental Genetics Group, University of Southampton. Areas of the heart that were investigated included the walls of the left atria and ventricle, the wall of the pulmonary artery and the pulmonary valve (see fig 5.3). An indirect detection method was used to visualise the protein under investigation. Once bound to its target the primary antibody was in turn bound by a secondary biotinylated antibody. The protein of interest was then visualised using steptavidin conjugated to horse radish peroxidase (HRP) which is detected enzmatically using 3,3' diaminobenezidine (DAB). Expression is ultimately detected as brown staining on the tissue section. Antibodies were optimised to identify the best conditions necessary for antigen unmasking using a combination of trypsinisation and boiling in sodium citrate (see appendix i). Appropriate concentrations required for primary, secondary and steptavadin can also be found in appendix i. Negative controls with no primary antibody were carried out in conjunction, and no background staining was seen.

The left atria (fig 5.4): Nidogen-1 was expressed strongly throughout the atrial myocardium. Expression was general and widespread, with no specific expression pattern as seen with collagen type IV and laminin-1. Relative to nidogen-1 and laminin-1, collagen type IV was faintly expressed within the atrial wall. Expression was seen as faint discrete lines. Strong expression for laminin-1 was detected. Similar to collagen type IV, laminin-1 was expressed as discrete lines throughout the atrial myocardium. These lines may demarcate the basement membranes.

The left ventricle (fig 5.5): Strong nidogen-1 expression was detected throughout the ventricular myocardium. The strongest expression was seen within the trabeculated muscle. Less nidogen-1 expression was seen in the outer compact myocardial layer. A strong line of nidogen-1 expression was seen in the thin exterior epicardial layer. Similar to the atrial wall, nidogen-1 was expressed in a ubiquitous manner with no discrete patterning. Collagen type IV was weakly expressed within the trabeculated layer. Little or no expression was detected in either the outer compact layer or the epicardium. Laminin-1 was strongly expressed within the trabeculated layer of the

ventricular myocardium. Similar to nidogen-1 reduced expression was seen in the outer compact layer, but with a strong line of expression in the epicardium. Unlike the atrial expression, laminin-1 was expression was widespread with no distinct line patterning.

The pulmonary valve (see fig 5.6): Nidogen-1 staining was detected within the cusps of the pulmonary valve. Expression was considerably weaker than previously seen within the atrial and ventricular walls. Interestingly, no expression was seen in the single-celled endothelial lining of the valves. Strong collagen type IV expression was detected within the valve cusps. No expression was detected within the endothelial layer. High levels of laminin-1 staining were seen in the cusps of the valve. Similar to nidogen-1 and collagen type IV, little expression was detected within the endothelial layer.

The pulmonary artery (see fig 5.7): Nidogen-1 staining was seen throughout the pulmonary artery, apart from the endothelial lining where expression was very much reduced or absent. In addition particularly strong circles of nidogen-1 staining were present within the pulmonary wall. Collagen type IV and laminin-1 were both expressed strongly within the pulmonary wall. Their expression appeared to be stronger than nidogen-1's in the pulmonary wall. Expression for both proteins was absent from the endothelial layer.

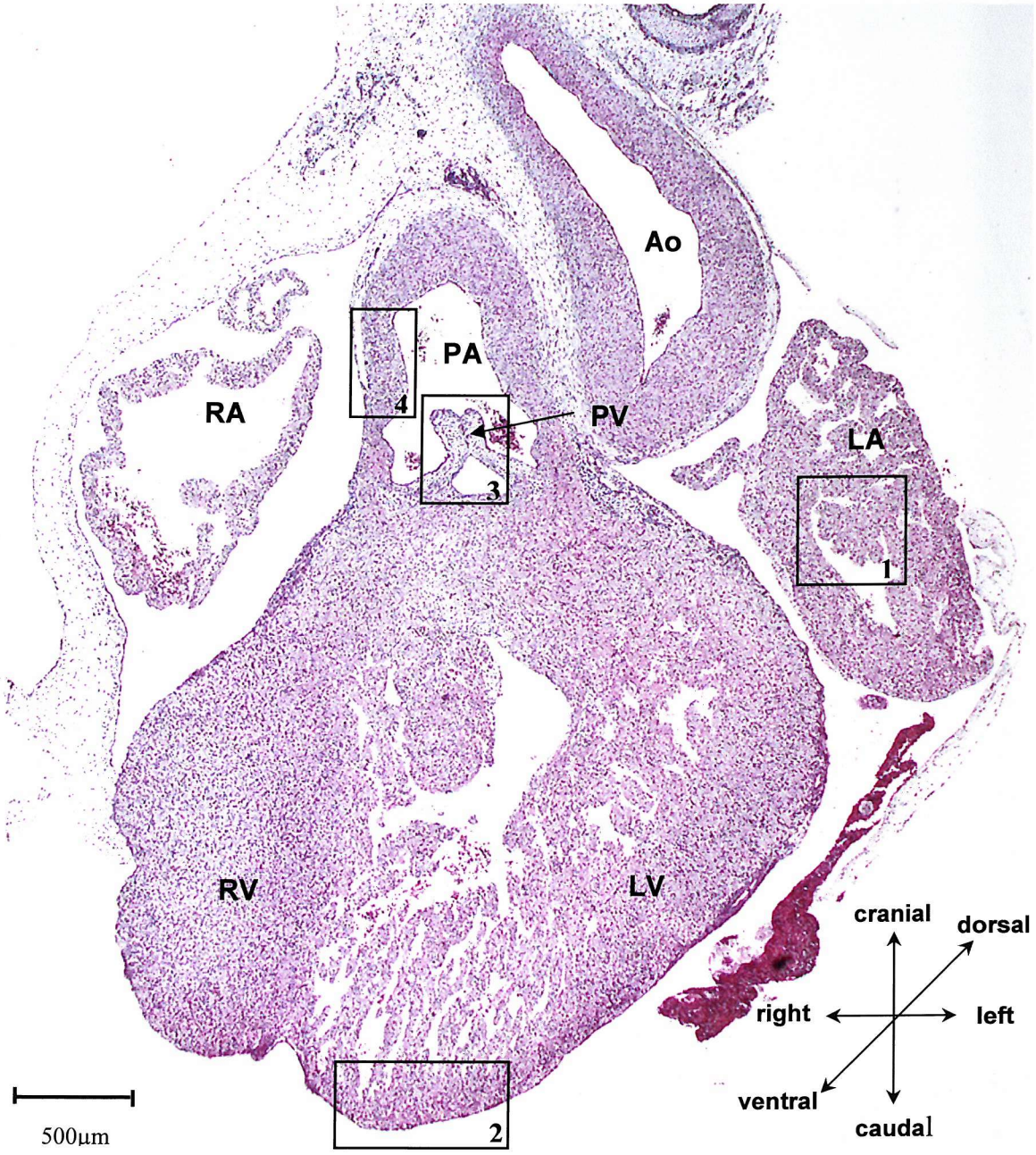

Fig 5.3: An H + E stained section of an eight week human fetal heart. The four boxed areas represent the four regions of analysis. 1 – The left atria; 2 – The left ventricle; 3 – the pulmonary valve; 4 – the pulmonary wall. RA – right atra; LA – left atria; LV – left ventricle; RV - right ventricle; Ao – aorta; PA – pulmonary artery.

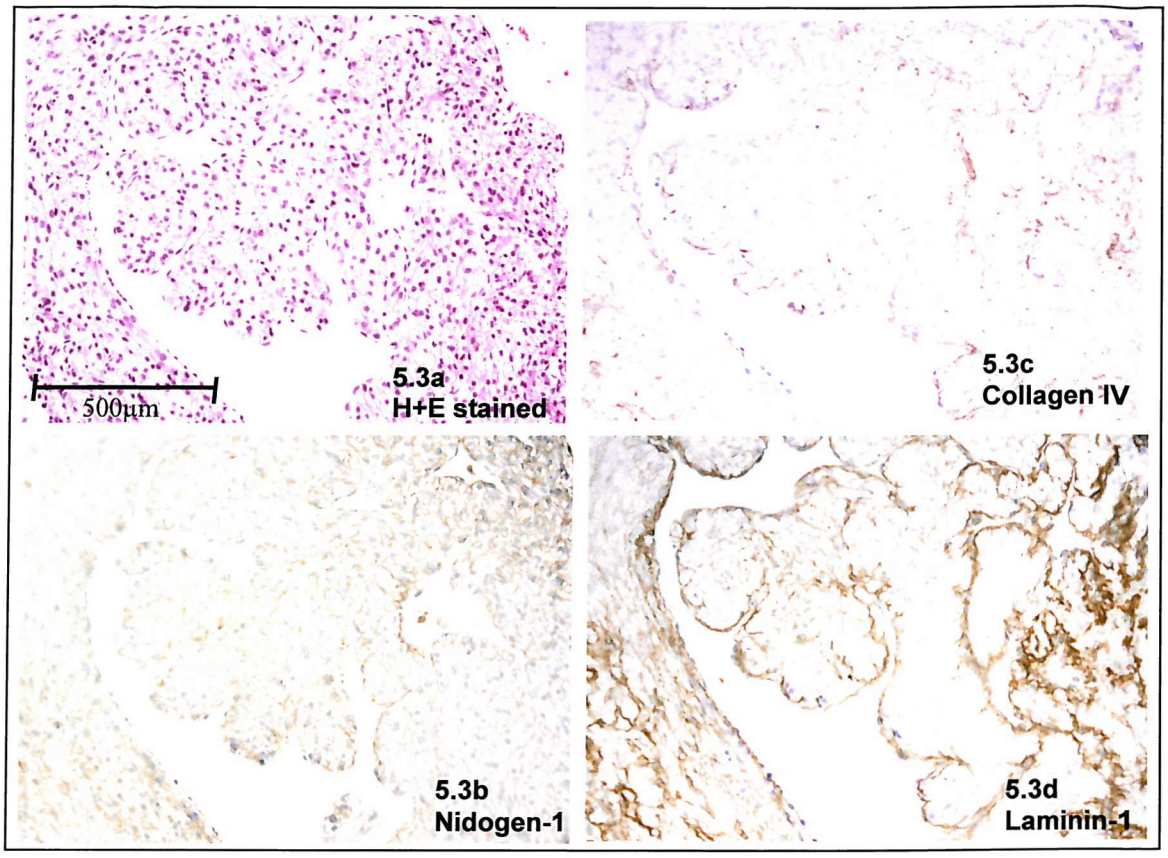

Fig 5.4: A figure showing the results of immunohistochemistry analysis of the left atria of an 8 week fetus. 5.4a: An image of an H+E stained section of part of the left atrial wall. 5.4b: The left atria investigated with a nidogen-1 antibody. Nidogen-1 was expressed strongly throughout the wall of the left atria. Expression was widespread with no distinct patterning as seen for collagen IV and laminin-1. 5.4c: The left atria investigated with a collagen type IV antibody. Weak collagen type IV expression was detected in the left atria. Unlike nidogen-1, collagen IV expression was seen as faint discrete lines throughout the atrial myocardium. 5.4d: The left atria investigated with a laminin-1 antibody. Laminin-1 was expressed strongly in the atrial wall. Similar to collagen type IV, laminin-1 was expressed as distinct lines running throughout the wall.

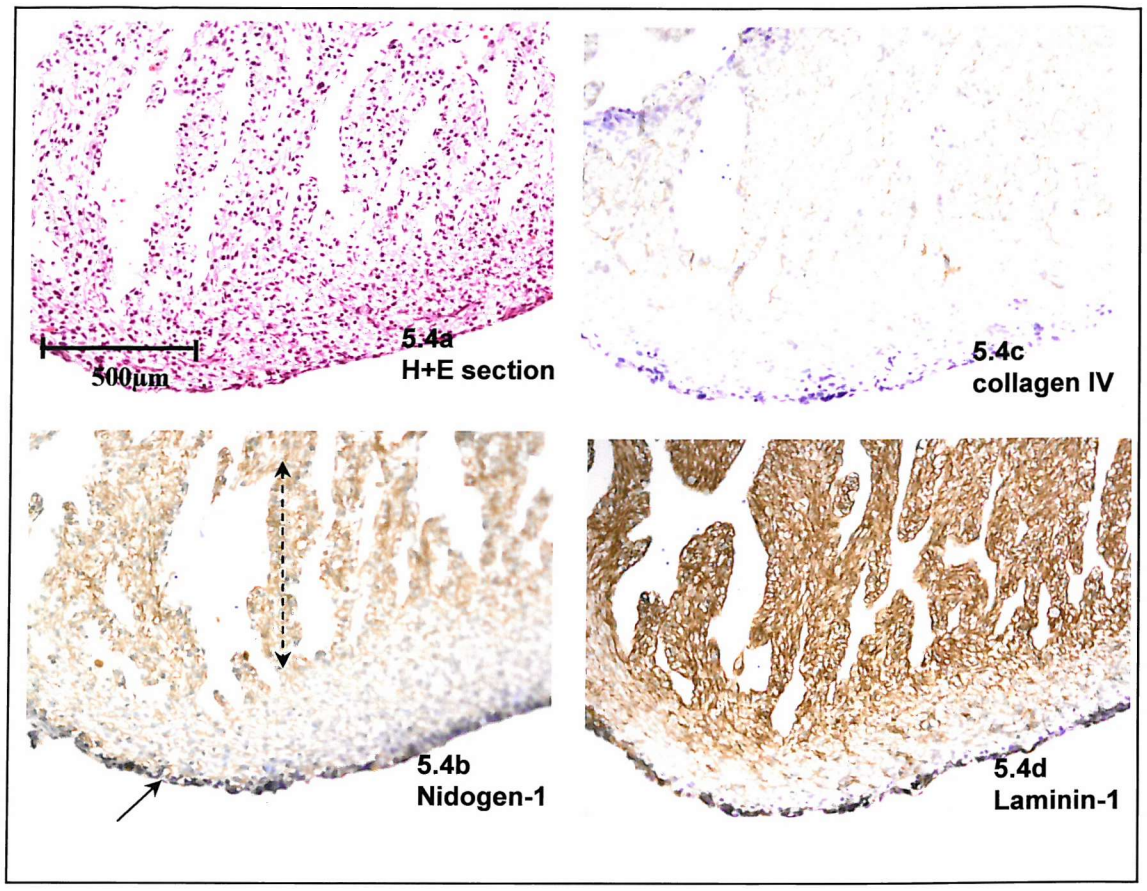

Fig 5.5: A figure showing the results of immunohistochemistry analysis of the left ventricular wall of an 8 week fetus. 5.5a: An image of an H+E stained section of part of the left ventricular wall. 5.5b: The left ventricle investigated with a nidogen-1 antibody. Strong nidogen-1 expression was detected throughout the ventricle wall. The strongest expression was seen in the trabeculated muscle (demarcated by the dashed double-headed arrow). Expression appeared to be reduced within the outer compact layer (located posterior to the trabeculated muscle). A strong line of expression was seen in the exterior of the wall, the epicardium (indicated by the arrow). 5.5c: The left ventricle investigated with a collagen type IV antibody. Faint collagen type IV expression was detected within the trabeculated muscle. Little or no expression was seen in the outer compact layer or the epicardium. 5.5d: The left ventricle probed with a laminin-1 antibody. Strong laminin-1 expression was detected within the trabeculated muscle of the ventricle. Similar to nidogen-1, reduced levels were detected in the outer compact myocardium. A strong line of expression was seen in the exterior epicardium.

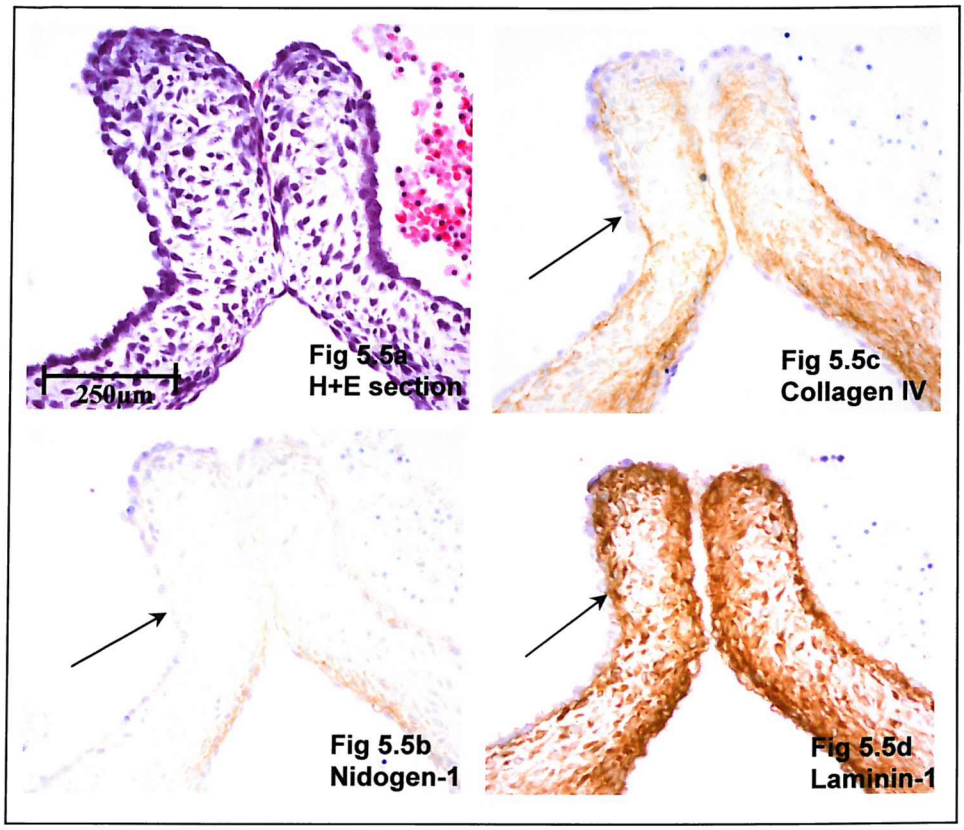

Fig 5.6: A figure showing the results of immunohistochemistry analysis of the pulmonary valve of an 8 week fetus. 5.6a: An image of an H+E stained section of the pulmonary valve. 5.6b: The pulmonary valve investigated with a nidogen-1 antibody. Nidogen-1 was detected within the cusps of the pulmonary valve. Expression was not as strong as detected within the atrial and ventricular walls. Expression was markedly absent from the endothelial lining of the valves (indicated by the arrow). 5.6c: The pulmonary valve investigated with a collagen type IV antibody. Strong levels of collagen type IV expression were detected in the valve cusps. Unlike in the atrial wall, expression was widespread and not present in a defined pattern. No expression was detected with the endothelial lining. 5.6d: The pulmonary valve investigated with a laminin-1 antibody. Similar to nidogen-1 and collagen type IV, laminin-1 was detected within the valve cups but was absent from the single-celled endothelial layer.

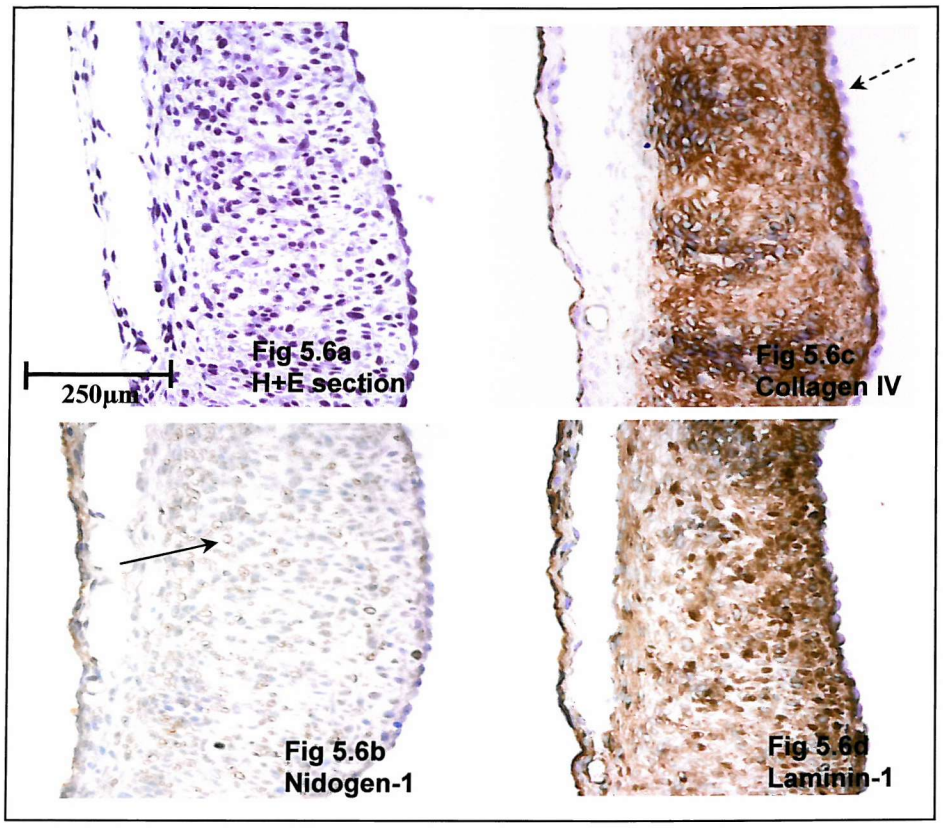

Fig 5.7: A figure showing the results of immunohistochemistry analysis of the pulmonary wall of an 8 week fetus. 5.7a: An image of an H+E stained section of part of the pulmonary wall. 5.7b: The pulmonary wall investigated with a nidogen-1 antibody. Strong levels of nidogen-1 staining were found in the wall of the pulmonary artery. At points throughout the wall, circles of strong nidogen-1 expression can be seen (as indicated by the arrow). These may represent blood vessels within the pulmonary wall. Similar to the pulmonary valve, reduced expression was seen in the endothelial lining of the artery. 5.7c: The pulmonary wall investigated with a collagen type IV antibody. High levels of collagen type IV protein were demonstrated within the pulmonary wall. Expression was absent from the endothelial lining (as indicated by the dashed arrow). 5.7d: The pulmonary wall investigated with a laminin-1 antibody. High levels of nidogen-1 expression were identified within the pulmonary wall. Expression appeared to be absent from the endothelial lining.

5.3.2 Mutation analysis of nidogen-1

RT-PCR and immunolocalisation demonstrated that nidogen-1, and the other basement membrane components, collagen type IV and laminin-1, are expressed in human fetal heart. This confirmed the candidacy of nidogen-1. Analysis of a panel of 20 patients was undertaken to identify mutations in the nidogen-1 gene (patients 12-31). All patients had outflow tract obstructive lesions and hypoplastic left heart. Amplification of all 20 nidogen-1 exons was possible using 25 PCR primer pairs (listed in appendix I). Primers were designed 50bp upstream and downstream of each exon to ensure complete sequence of the intron/exon boundary could be obtained. Heterozygous changes were identified initially by dHPLC analysis. Sequence variations were represented by dual peaks on the dHPLC trace. Direct sequencing was subsequently carried out to identify the nature of each change. A panel of 60 DNA samples from normal controls were screened for identified changes to determine whether they were polymorphisms.

The results are summarised in table 5.1. 10 sequence changes were identified. These are listed below and are highlighted in nidogen-1 sequence in appendix iii.

Exon 2: A heterozygous cytosine to thymine substitution (489C>T) found in patients 15, 16, 22, 24, 25 and 26. No amino acid change was predicted. This was a known SNP (RS6665008; http://www.ncbi.nlm.nih.gov/SNP/) and was identified in two normal control samples by sequence analysis.

Exon 8: A heterozygous guanine to adenine substitution (1914G>A) found in patients 22 and 28. No amino acid change was predicted. This was not a known polymorphism but and was identified in two normal control samples by sequence analysis.

Exon 9 (see fig 5.8): A heterozygous guanine to adenine substitution (2061G>A) found in patient 28. No amino acid change was predicted. This was not a known polymorphism but was identified in a normal control sample by sequence analysis.

Intron 9: A heterozygous adenine to guanine change (IVS9+5A>G) identified in patients 18, 26 and 30. The substitution was located five base pairs downstream of

exon 9. This change was identified in two normal control samples by sequence analysis.

Exon 12: An adenine to guanine substitution (2272A>G) identified in patients 15 and 16. No amino acid change was predicted. This was a known polymorphism (RS2031487) and was identified in two normal control samples by sequence analysis.

Exon 13: An adenine to guanine substitution (2444A>G) identified in patients 24 and 25. No amino acid change was predicted. This was not a known polymorphism and was not identified in 120 chromosomes by dHPLC analysis.

Intron 13: A cytosine to adenine substitution (IVS13+2579C>A) identified in patients 14 and 22. The variation was located 7bp upstream of the exon start. This was not a known SNP but and was identified in two normal control samples by sequence analysis.

Exon 16 (see fig 5.9): A thymine to cytosine substitution (3055T>C) identified in patients 14, 22 and 27. This was a known polymorphism (RS3738525) and was confirmed in 2 control samples by sequence analysis.

Intron 18 (see fig 5.10): A homozygous guanine to adenine substitution (IVS18+710G>A + IVS18+710G>A) identified in patient 14. This change was located 4bp upstream of the exon start site. This was not a known polymorphism. It was identified in heterozygous form in 1 out of 70 control chromosomes. No homozygous changes were identified.

Exon 20: An adenine to guanine substitution (3605A>G) identified in patients 14, 22 and 27. This resulted in a glutamine to arginine amino acid change. This was a known polymorphism (RS3213190) and was found in 2 control samples by sequence analysis.

Exon/Intron	Patient number	DNA change	Protein change	Comments
Exon 2	pt 15, 16, 22, 24, 25, 26	489C>T	No predicted change	Known polymorphism (RS6665008); confirmed in two out of forty control chromosomes
Exon 8	pt 22, 28	1914G>A	No predicted change	Confirmed in two out of forty control chromosomes
Exon 9	pt 28	2061G>A	No predicted change	Confirmed in one out of forty control chromosomes
Intron 9	pt 18, 26, 30	IVS9+5A>G	-	Confirmed in two out of forty control chromosomes
Exon 12	pt 15, 16	2272A>G	No predicted change	Known polymorphism (RS2031487); confirmed in two out of forty control chromosomes
Exon 13	pt 24,25	2444A>G	No predicted change	Not identified in 120 chromosomes by dHPLC
Intron 13	pt 14, 22	IVS13+2579C>A	-	Confirmed in two out of forty control chromosomes
Exon 16	pt 14, 22, 27	3055T>C	No predicted change	Known polymorphism (RS3738525); confirmed in two out of forty control chromosomes
Intron 18	pt 14	IVS18+710G>A	-	Identified as heterozygous change in 1 out of 70 chromosomes. No homozygotes found
Exon 20	pt 14, 22, 27,	3605A>G	Glutamine - arginine	Known polymorphism (RS3213190); confirmed in two out of forty control chromosomes

Table 5.1: A table summarising the results of sequencing analysis. A panel of 20 CHD patients were analysed for sequence changes within the nidogen-1 gene. 10 DNA changes were identified by dHPLC and sequencing analysis. Only one in exon 20 resulted in an amino acid change from a glutamine residue to arginine. Only two variants, those located in exon 13 and in intron 18, were either not a known SNP or were not found in control chromosomes. The change in exon 13 resulted in no predicted amino acid change, but was not identified in 120 control chromosomes by dHPLC analysis. The variation in intron 18 was a homozygous change and this was identified only in heterozygous in 70 control chromosomes by sequencing analysis. No homozygous changes were identified in the control DNA panel.

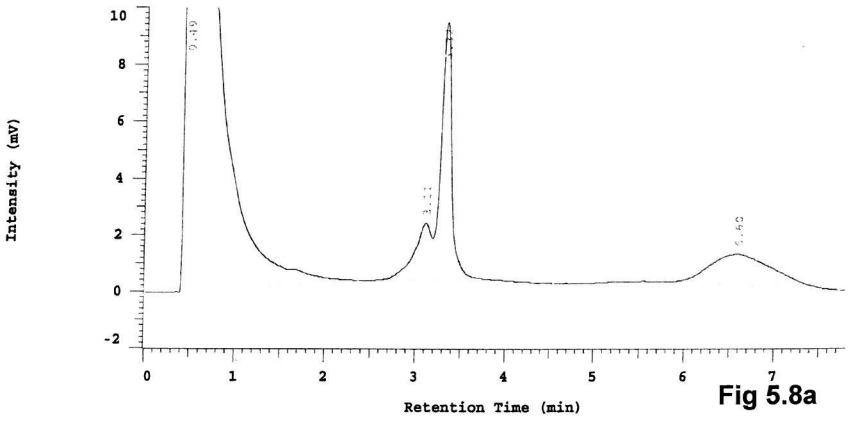

Fig 5.8a

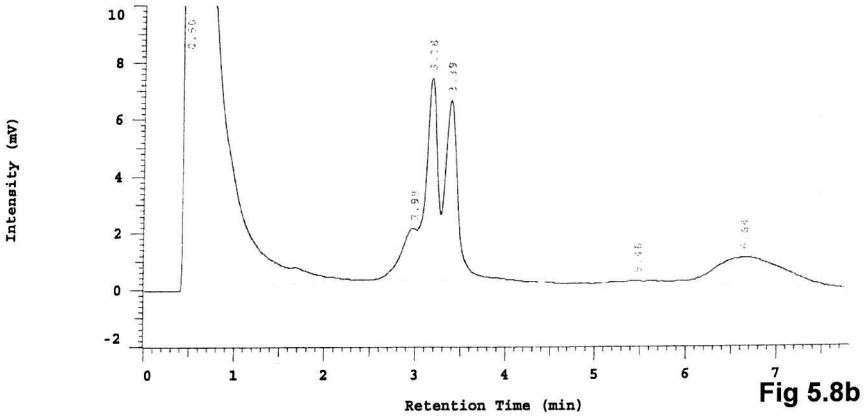

Fig 5.8b

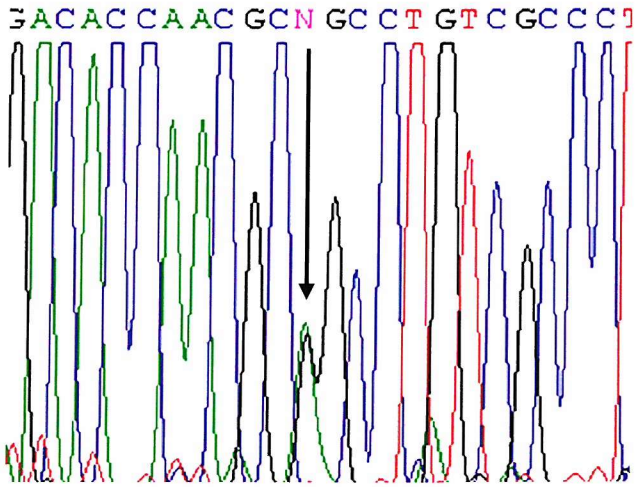

Fig 5.8c

Fig 5.8: The results of mutation analysis of exon 9 carried on patient 28. 5.8a dHPLC trace of normal control DNA for exon 9. A single peak indicated that there were no sequence mismatches 5.8b dHPLC trace of patient 28 for exon 9. A dual peak suggested the presence of a sequence mismatch within exon 9 of this patient. 5.8c Sequencing analysis of patient 28 for exon 9 revealed a guanine to adenine substitution, as indicated by the arrow. This is represented by the letter 'N' in the sequence found at the top of the figure.

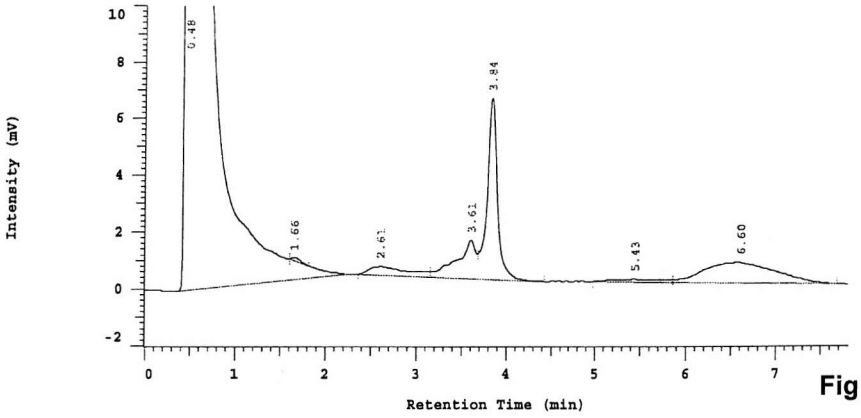

Fig 5.9a

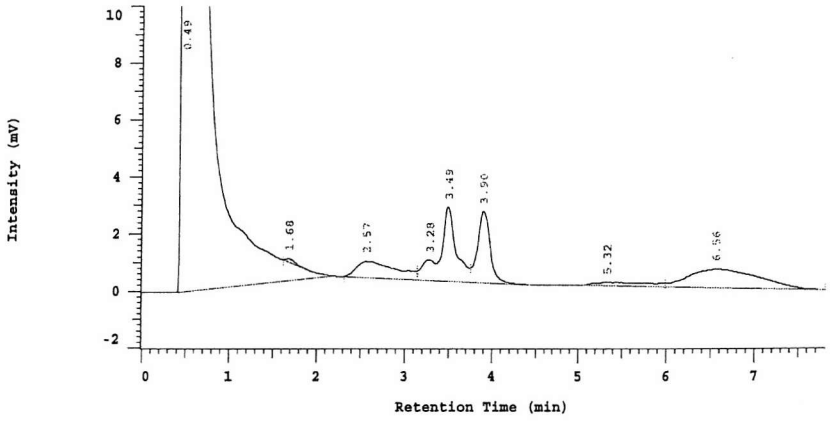

Fig 5.9b

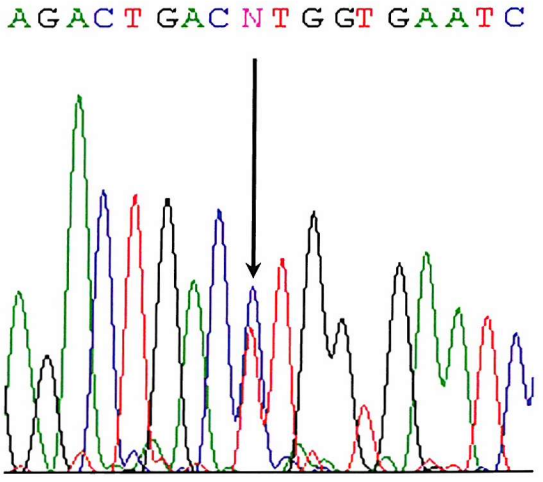

Fig 5.9c

Fig 5.9: The results of mutation analysis of exon 16 carried out on patient 22. 5.9a dHPLC trace of exon 16 for a normal control DNA sample. 5.9b dHPLC trace of exon 16 for patient 22. 5.9c Sequence trace of exon 16 for patient 22. A thymine to cytosine substitution was identified (as indicated by the arrow)

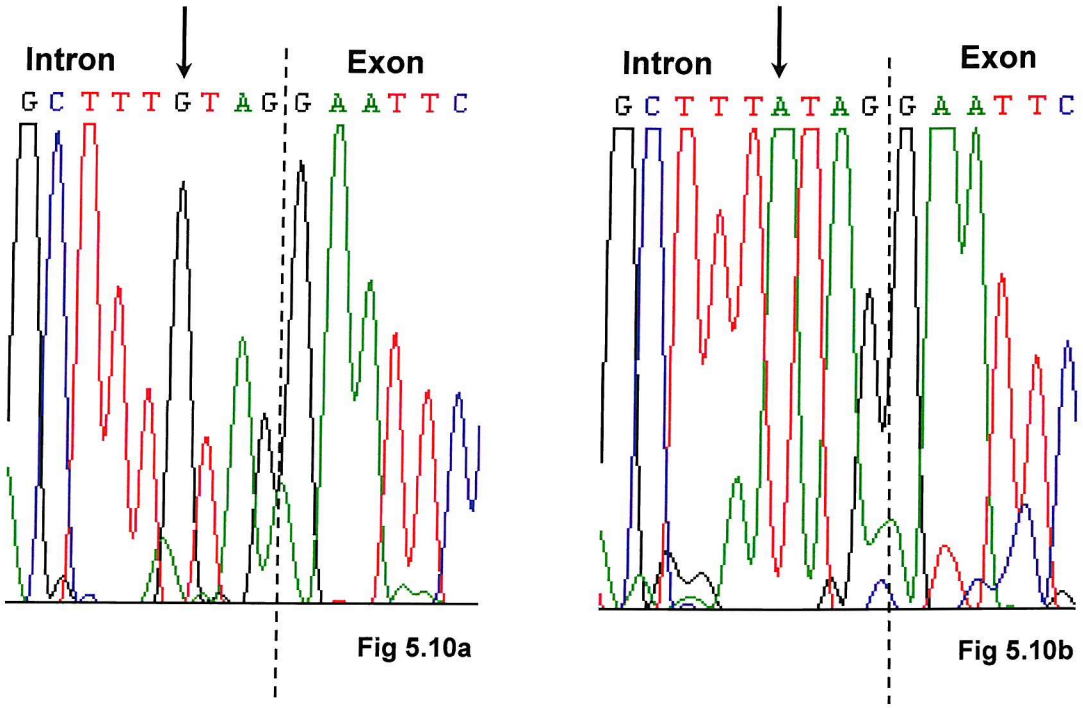

Fig 5.10: The results of sequencing analysis of intron 18 carried out on patient 14. 5.10a Sequence of intron 18/exon 19 boundary for a normal control DNA. The dotted line represents the exon start site. 5.10b Sequence of intron 18/exon 19 boundary for patient 14. A homozygous guanine to adenine substitution was identified in this patient. This change was identified only in heterozygous form in one chromosome out of 70 sequenced.

5.4 Discussion

Within the 9Mb proximal critical region defined by patient 1, the basement membrane protein, nidogen-1, was identified as a strong candidate gene. The basement membrane is thought to play an important role in development. Gene knock-out studies have shown that components of the BM are crucial for cardiac development. A hypothesis can be put forward which suggests that components of the extracellular matrix, such as the elastic fibres and the endocardial cushions, are anchored to the underlying basement membrane through interaction with nidogen-1. This may take place through intermediaries, in particular members of the fibulin family. Mutations in nidogen-1 may disrupt these developmental pathways, resulting in obstructive outflow tract defects due to disruption to the elastic fibres and sepation abnormalities because of aberrant endocardial cushion development. However the nidogen-1 null mouse displayed no observed pathology, this may be due to genetic redundancy with nidogen-2 being up-regulated. There exist however important differences in the binding potentials of nidogen-1 and nidogen-2 between human and mouse.

Expression of nidogen-1 within human fetal heart was demonstrated using RT-PCR and immunohistochemistry. Immunolocalisation experiments were also carried out for the two main basement membrane components, laminin-1 and collagen type IV. Interestingly, although these proteins are expected to interact within the basement membrane, differences in their expression patterns become apparent. Within the atria, nidogen-1 had a ubiquitous expression pattern. Both collagen type IV, which was weakly expressed, and laminin-1 were expressed in definite, discrete lines, which may represent the basement membrane. The more widespread staining seen for nidogen-1 may suggest that its role within the atrial myocardium is not restricted to the basement membrane. Differences were also apparent within the ventricular wall. For both nidogen-1 and laminin-1 high levels of expression were demonstrated with the trabeculated muscle, reduced expression was seen in the outer compact layer. A strong band of staining for these two proteins was also seen around the exterior of the ventricular myocardium, the epicardium. Collagen type IV was restricted to the trabeculated muscle. No expression was seen in the outer myocardium of the epicardium. All three proteins were expressed within the pulmonary valve. Interestingly no expression for the three proteins was seen within the single-celled

endothelium. Similarly within the walls of the pulmonary artery, nidogen-1, collagen type IV and laminin-1 were all strongly expressed, with little or no staining detected within the endothelial layer. It is intriguing to note the circular rings of strong nidogen-1 expression within the arterial wall. One explanation may be that these demarcate smaller blood vessels within the wall. Alternatively these could be cells secreting nidogen-1. The high levels of nidogen-1 expression within the myocardium of the pulmonary wall and the ventricular myocardium is an expression pattern that fits well with the outflow tract defects and ventricular septal abnormalities seen in 1q deletions patients.

Mutation screening of a panel of 20 CHD patients was subsequently undertaken. Ten nucleotide changes were identified. Seven of these were in exonic sequence. Only one, 3605A>G, resulted in an amino acid change from a glutamine to an arginine residue. Given that this results in a change from an uncharged polar to a charged polar amino acid it is possible that this could have an impact on protein folding and structure. This however was a known SNP (RS3213190), and was also identified in control DNA samples, suggesting that it did not have any affect on function. The six remaining exonic changes were predicted to have no affect on amino acid coding. It is possible that 'silent' nucleotide changes that give rise to no amino acid alterations can affect gene funtion through interruption of consensus sequeneces required for exon splicing, or the introduction of ectopic splice sites (section 1.4.3; reviewed in Cartegni et al., 2002). However only one of the changes, 244A>G, was not a known SNP, which would suggest that they are not pathogenic. In addition all five were identified in control chromosomes providing further confirmation that they were not pathogenic.The nucleotide change found in exon 13, 2444A>G, which was not a known SNP, was also not identified in 120 control chromosomes by dHPLC. Located 50 bp downstream of the exon start site it is unlikely to affect exon splicing. As it also does not introduce any amino acid change, at this stage it seems likely that this change is a rare polymorphism, rather than a pathogenic mutation. Unfortunately parental samples were unavailable to show whether this was a *de novo* change. Of the three intronic changes, two were identified in the control DNA samples. The first was an adenosine to guanine change identified in intron 9, 5bp downstream of exon 9 (IVS9+5A>G). The second was within intron 13, located 7bp upstream of exon 14 (IVS13+2579C>A). As these were both found in control samples, they are unlikely to

be of importance. The last change identified was a homozygous change (G-A) located 4 bp up from the start of exon 19 in patient 14 (IVS18+710G>A + IVS18+710G>A). The exon had been sequenced because of a suspected heteroduplex peak. No heterozygous change was found. Investigation of control samples identified the change in heterozygous form in one out of seventy control chromosomes analysed. Parental samples were not available for analysis. It is possible that this may interrupt the exon splice site. Splicing is carried out by the splicing complex, the spliceosome, which has two primary functions. Firstly, to recognise and bind the exon/intron boundary, and secondly to remove the introns and re-join the exons. The identification of the exon/intron boundaries occurs through three recognition consensus sequences, the 5' splice site, the 3' splice site, and the branch site (see fig 5.11). The 3' splice site is preceded by a polypyridimine tract of variable length. The branch point is typically located 18-40 nucleotides upstream of the polypyridimine tract. The fourth base pair upstream of the exon start site can be either of the four nitrogenous bases (see fig 5.11). It therefore may be unlikely that the change identified in intron 13 of patient 14 is of significance. However whilst this base pair position can vary between genes, correct splicing of exon 18 of nidogen-1 may require that this position remains invariant to allow efficient recognition of the splice site. Given that this was identified in homozygous form, it would be expected that, if benign, this would be a common polymorphism. Interrogation of the control panel suggests that this is not the case, with the change only being found in heterozygous form in 1 out of seventy chromosomes sequenced. It is difficult to ascertain at this stage its functional significance. It would be extremely interesting to determine whether this change does have any influence on exon splicing and consequently on protein production (see future work – section 6.2.2). Although no obvious pathogenic mutations have been found this does not exclude nidogen-1 as a CHD candidate gene.

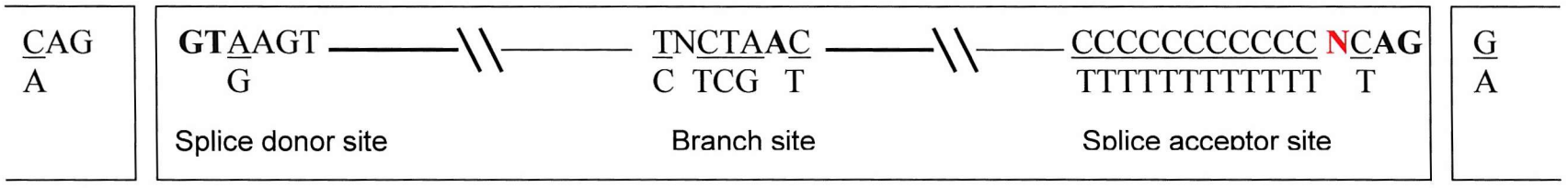

Fig 5.11: A diagram showing the consensus sequences required for exon splicing. The recognition sequences required for splicing are the 5' donor site, the branch site and the 3' acceptor site. Nucleotides highlighted in bold are in most cases invariant. Other nucleotides represent the majority nucleotide found at this particular position. A guanine to adenine substitution was identified in intron 18 of patient 14 four base pair upstream of the exon start (highlighted in red).

Chapter 6: Summary and future directions

6.1 Summary

Chromosomal deletions within locus 1q42-44 can be a cause of congenital heart disease. A significant association with ventricular septal defects has been demonstrated (Brewer et al., 1998). More severe forms of cardiac pathology, such as hypoplastic left heart, have been documented (Wright et al., 1986). It is therefore likely that within this chromosomal region is a crucial CHD disease gene, which must be present in the correct dosage to allow cardiac development to proceed correctly. Little molecular data on these 1q42-44 patients was available within the literature. Consequently, it was difficult to establish clear critical disease regions to effectively enable the identification of CHD susceptibility genes. The initial aim of this thesis was to define a molecular CHD critical region within 1q42-44 to identify candidate genes for investigation

Molecular analysis was carried out on a set of 10 patients with known 1q deletions; seven of these had a cardiac defect. The data obtained suggested that there may be two CHD susceptibility regions; one located proximally within 1q42.3 and the other terminally within 1q44. However at this stage it cannot be discounted that a CHD gene lies in between these two regions which is disrupted through position effect when either region is deleted. To attempt to reduce the critical regions a panel of 15 patients with normal karyotypes, was investigated for submicroscopic deletions, using a combination of loss-of heterozygosity and fluorescent *in situ* hybridisation analysis. No deletions were identified.

The need for FISH confirmation limited the number of patients that could be investigated, as immortalised cell lines were used to obtain metaphase chromosomes. In addition FISH is a time-consuming methodology with restricted through-put. A microarray-CGH approach would increase the number of patient samples that could be investigated, improve analysis throughput and would permit the analysis of several genomic regions to be carried out in parallel. Previous studies had demonstrated the utility and reliability of microarray-CGH for the identification of chromosomal imbalances (Pinkel et al., 1998). A proof-of-principle study was therefore carried out to develop array slides to show the feasibility of this approach for detecting deletions within 1q42-44. Microarray-slides were developed with genomic coverage of the

proximal 1q deletion region defined from analysis of patient 1. Clones from the 11q HLH critical region were also arrayed (Phillips et al., 2002). Preparation of the slides was carried out using a modified DOP-PCR method as previously described (Fiegler et al., 2003a). Hybridisation of labelled patient 1 DNA against a normal control confirmed that the slides could detect deletions. The deletion data obtained matched that previously obtained from microsatellite and FISH analysis confirming the reliability of the result. Having confirmed the reliability of the technique it is hoped that these proof-of-principle experiments will serve as a platform for future experiments, not only with regard to the investigation of 1q monsomy, but also other regions through-out the genome, deletions within which can cause CHD (see future work 6.2.1). It also demonstrated that although substantial time is required for the initial phase of array design and development, once established array-CGH will have a significant impact on through-put with what was the equivalent of several hundred FISH experiments being carried out along chromosome 1 in a single assay.

The majority of genes within the 1q42-44 locus were uncharacterised, especially within 1q44. Assessment of candidacy in all cases was therefore limited. Several lines of evidence supported the investigation of the basement membrane protein nidogen-1. Firstly basement membrane components are known to have important roles in tissue development and growth. Secondly it has potential interactions with components of the elastic fibre and the endocardial cushions. A plausible hypothesis can therefore be put forward suggesting that components of the elastic fibre, in particular fibrillin-1 and elastin, and the endocardial cushions are linked to the basement membrane through an interaction with nidogen-1 and members of the fibulin family. Haploinsufficiency of nidogen-1 may lead to disruption of elastic fibre development, possibly resulting in outflow tract stenotic lesions or abnormal endocardial cushion development, leading to septation defects. RT-PCR and immunohistochemistry analysis confirmed strong levels of nidogen-1 expression within human fetal heart. Mutation screening of nidogen-1 in a panel of 20 CHD patients identified two sequence changes, which were not known polymorphisms and which were not found in control samples. One was located within exon 13. It resulted in no amino acid change and was not thought to disrupt important regulatory sequences or introduce ectopic splice sites, and was therefore unlikely to be pathogenic. The other was an intronic homozygous change that was located 4bp upstream of the exon start. The

splice site is an important regulatory site, any variations in sequence may disrupt correct gene splicing. This change was only identified in heterozygous form in one control DNA sample out of 34. It is difficult at this stage to assess the functional significance of this base pair substitution. It would be interesting to ascertain what affect, if any, this sequence variation has on nidogen-1 splicing. Although no obvious disease causing changes have been identified, the uncertainty over the affect of this splice site change and the fact that only a panel of 20 patients were interrogated suggests that nidogen-1 can not at this stage be ruled out as a candidate gene for congenital heart disease.

6.2 Future directions

6.2.1 Microarray technology

Reducing the critical regions remains a priority for future research. Microarray-CGH may be the most efficient method for achieving this. The DOP-PCR amplified target DNA from the 1q critical region, the Sanger Institute chromosome 1 clone set, and the 11q HLH region are available for the printing of further slides. Work is now on-going to ensure full coverage of both critical regions, as more complete sequence is now available, and to extend coverage distally to the terminal 1q disease region. A 'cardiac array' is under development, which, in addition to 1q42-44 and 11q25, will allow investigation of several other genomic regions for loss or gain of chromosomal material that are known to cause CHD.

Microarray-CGH is likely to have a significant impact in the diagnosis of genetic disease and on deletion mapping studies to identify disease genes. A recent study demonstrated the development of arrays with complete genome coverage, consisting of 32,433 overlapping BACs being printed on to a single slide (Ishkanian et al., 2004). Verification of these arrays demonstrated the identification of a 300kb microamplification within chromosome 13q and a 240kb microdeletion within 7q22.3 in a breast cancer cell line. It is possible that in the foreseeable future microarray-CGH will replace karyotyping analysis, which is currently used for genome-wide genetic diagnosis. This may increase the chance that smaller microdeletions be identified which would facilitate the identification of a cardiac disease gene within the

1q42-44 region. The utility of array-CGH in disease gene identification was demonstrated in a recent investigation of CHARGE syndrome (OMIM: 214800; (Vissers et al., 2004)). CHARGE syndrome is a common cause of developmental abnormalities with an estimated birth incidence of 1:12,000. Cardiac defects can be a feature. Vissers *et al.* optimised a genome-wide BAC array with 1-Mb resolution and identified an approximate 5 Mb deletion within 8q12 in one patient. An array of 918 overlapping BAC clones covering chromosome 8 was subsequently established to permit investigation of additional patients for abnormalities of chromosome 8. A patient with a known balanced chromosome 8 translocation was shown to have two microdeletions that overlapped with that of the original deletion patient. A critical disease region of 2.3Mb within 8q12 was defined. Sequencing of the coding regions and intron-exon boundaries of the nine known or predicted genes was carried out. Ten heterozygous mutations were identified in the *CHD7* gene, seven of which introduced a premature stop codon. One was a *de novo* change located 7bp upstream of exon 23, which was could possible have affected exon splicing. The final two were missense mutations. Recent studies have also demonstrated that microarray technology can be used to detect chromosomal translocations through hybridising flow-sorted chromosomal material on to DNA microarrays in a process termed 'array painting' (Fiegler et al., 2003b; Gribble et al., 2004).

Current microarray-CGH protocols predominantly use DNA propagated from genomic clones as hybridisation targets. The use of BAC and PAC clones however has important limitations, leading to the recent development of alternative strategies for array construction, which may have bearing on the investigation on 1q42-44 monosomy in the future. The first limitation is that the resolution of analysis is limited by the size of the insert. Second pre-selection of arrayed sequence is not possible. This results in an array, which may contain a high number of common repeats and sequence similarity via pseudogenes, or paralogs. This may lead to increased background fluorescence, which was a possible reason for the result seen for clone BA528D17 (see section 4.3), and which has also been seen in other studies (Mantripragada et al., 2004). Third, the correct clone is not always delivered and there is the potential of bacterial phage contamination. These limitations have lead to the development of arrays in which all commonly repeated and redundant sequences were excluded from the array. In 2003, Mantripragada *et al.*, designed coding sequence

specific, free-repeat arrays for detecting deletions within the neurofibromatosis type 2 gene *(NF2)* to allow greater resolution of deletion detection over a genomic clone based approach (Mantripragada et al., 2003). Unique sequences form the NF2 gene were PCR amplified from cosmids covering the locus. Validation confirmed the ability of these arrays to detect deletions within 3 *NF2* patients. The resolution of deletion detection is in the range of 15-20kb. An additional advantage was the reduction or complete removal of Cot-1 DNA. Not only is this an expensive reagent, but it has also been commented that variations in quality and content between batches has lead to unreliability in blocking repetitive sequence elements (Carter et al., 2002; Mantripragada et al., 2003; Mantripragada et al., 2004). This *NF2* gene-specific array is the first high-resolution tool for the detection of diagnostically significant gene copy number aberrations. FISH is the current method used to detect *NF2* deletions, however this may be replaced by array-based methodologies in the future.

This approach has recently been adopted for arrays targeting the DiGeorge region on chromosome 22q11 (Mantripragada et al., 2004). The 3Mb deletion region contains a high number of common repeats which lead to difficulty in analysis using a full coverage chromosome 22q array, with DNA extracted from genomic clones (Buckley et al., 2002). To enhance diagnosis of chromosomal imbalance with the 22q11 region, Mantripragada *et al.* sought to compare the use of genomic clones, repeat-free pools of amplified genomic DNA and cDNAs as hybridisation targets. On analysis of DiGeorge patients, considerable variation was seen in the results obtained from the genomic clones, which may be explained by the presence of high numbers of repeat elements. No reliable data could be obtained from using single cDNA clones or when only two cDNA clones were pooled. However when three or more cDNAs were pooled it was possible to consistently distinguish between one and two gene copies. Reliable data was consistently obtained from three repeat-free PCR amplified DNA fragments following analysis of six DGS patients. It was reported that as little as 11.5kb non-redundant, repeat free PCR-generated sequence can be used for the reliable detection of hemizygous deletions. In the case of cDNAs the results suggested that 3.5kb is sufficient for accurate detection of dosage alterations. These represent significant improvements on resolution from genomic clone arrays. For regions rich in redundant sequences and repeats a specifically tailored array-CGH approach may yield the most accurate results for gene copy number profiling. However it seems

unlikely that such specific, tailored arrays need to be applied to the investigation of large genomic regions, such as the critical regions within 1q42-44. Should these critical regions been significantly reduced in the future, the use of such high-resolution, gene-specific may be more appropriate.

Microarray hybridisation technology has also the potential to be used for mutation detection (Hacia, 1999). It is possible to screen a gene or DNA sequence of interest for base pair mismatches through the design of arrays, which contain olignucloetides matching all wild-type and single nucleotide substitutions in that stretch of DNA. Test DNA is PCR amplified, fluorescently labelled, and hybridised to the array in competition with a wild-type control sequence, labelled in a different colour. Measurement of the fluorescent ratios will indicate the presence of a mismatch. In 1996, Hacia JG *et al.* designed high-density arrays consisting of over 96,000 oligonucleotides 20bp in length to screen for a wide range of heterozygous mutations in the 3.45 kb of the heredity breast and ovarian cancer gene BRACI (Hacia et al., 1996). Fourteen out of fifteen patient samples with known mutations were accurately diagnosed. 8 SNPs were also detected. A similar approach was adopted recently to detect mutations in human mitochondrial DNA which are known to be a cause of some cancers (Maitra et al., 2004). Validation of this oligonucleotide array demonstrated that the array is able to sequence greater than 29kb of double-stranded DNA in a single assay with an overall frequency of 96% for successful sequence fidelity. Therefore there may exist the potential to investigate numerous candidate genes in parallel. Microarray technology is likely to have a significant role in mutation detection in the future both within diagnostics and in a candidate gene approach in understanding disease aetiology.

6.2.2 Identifying the disease gene and further investigations.

Improved sequence and greater characterisation of genes mapping to the critical regions will enable candidate genes within the 1q42-44 region to be targeted with greater efficiency. The majority of genes within 1q44 at the time of investigation were uncharacterised, and of those that were, none were obvious candidate genes. This situation has since changed. TGF(beta)-induced transcription factor 2 is now known to locate within 1q44 (see appendix ii). TGF(beta)-2, which maps to 1q41, is an

important developmental gene that is thought to play a crucial role within the endocardial cushions during septal and valve development (see section 1.2.2.5). Knock-out studies in mice of TGF(beta)-2 gave rise to an array of cardiac defects, which inlcuded double-outlet right ventricle, VSDs, pulmonary valve abnormalities, and malformations of the AV canal (Bartram et al., 2001). TGF(beta)-induced transcription factor 2 may function downstream of TGF(beta)-2. Haploinsufficiency of TGF(beta)-induced transcription factor 2 could therefore be a strong candidate gene for the the CHD seen in patients with terminal 1q deletions.

Although none of the changes identified within nidogen-1 were obviously pathogenic, it would be of interest to pursue the homozygous change identified within the exon 19 splice site found in patient 14. Investigation of a larger control panel may give more of an indication of the frequency of the change within the population. In addition it would be useful to obtain parental DNA to ascertain whether this was a *de novo* change. Should this be a rare variation within the population, and a novel change to the patient, further investigations would be justified to determine whether this sequence alteration does have any affect on gene splicing, or protein production or function. If nidogen-1 was shown to be a cause of CHD through the finding of pathogenic mutations, other downstream ligands, such as fibulin-2, which may also play important roles in the same cardiac developmental pathways, would be legitimate targets for investigation. Fibulin-2 has been identified as a CHD candidate gene in the investigation of 3p- syndrome, but was subsequently discounted as it was found to map outside the deletion region (Green et al., 2000). The identification of novel interactions would be of interest. Nidogen-1 may interact with the elastin microfibril interface located proteins, emilin-1 and emilin-2. Emilin-1 is found at the interface between amorphorous elastin and the fibrillin associated microfibrils in the elastic fibre, similar to fibulin-2 (Bressan et al., 1993). It is thought that it might regulate the formation of the elastic fibre given the finding that elastin deposition *in vitro* is perturbed by the addition of anti-Emilin antibodies (Bressan et al., 1993). The recent knock-out demonstrated defects in elastic fibre development in the aorta and the skin (Zanetti et al., 2004). A second emilin gene, emilin-2, has been identified which was shown to be specifically expressed during fetal development within the heart by RNA blot hybridisation (Doliana et al., 2001). RT-PCR carried out during the course of this

thesis confirmed this finding (data not shown). Emilin-2 maps to chromosome 18p; deletions of which can be a cause of CHD.

The identification of the CHD disease-causing gene within the 1q42-44 region would have significant implications for clinical practice, in particular with regard to genetic counselling. Recurrence risk could be predicted with greater certainty enabling more effective counselling be given to existing parents or couples considering a family, one of whom may have a cardiac defect. It may also offer the possibility of antenatal investigation, and give the option of pregnancy termination should that be desired. Although such investigations should be treated with caution due to the variability in penetrance associated with many potentially disease causing mutations. Advances in paediatric surgery have meant that many previously life-threatening conditions now have much improved survival rates. Future parents could therefore be given more full information regarding risk and the outcomes of surgical intervention, allowing them the make more fully informed decisions. The identification of a 1q CHD gene would enable investigations to determine whether any environmental factors modify disease severity, and if so, if changes in parental lifestyle could improve potential outcomes. In addition some patients may wish to simply know what the cause behind their condition is, and the provision of a simple diagnostic test would be of great benefit to these patients.

Appendix i

PCR primers and antibodies used in this thesis

Primers used for microsatellite analysis

Primers were obtained from Qiagen Operon, Crawley, UK. Primers with light sensitive modifications were stored in darkened eppendorfs.

Primer	Primer sequence	Modification	Amplicon size (bp)	PCR temp. (°C)
D1S2880F	CGTGGTTCTAATCGG	FAM	109-135	56
D1S2880R	CATCATTTGCTTGCTGC			
D1S2871F	TGAAGTGTGCATTCTTACAT	TET	215-241	56
D1S2871R	CGAGACATTTGCATCATC			
D1S2763F	ACTTTTGAATGTTTTCGG	TET	161-173	54
D1S2763R	GAGACAGAGTTTTACCATCAT			
D1S439F	CACAGACTTCATTAGAGGGG	HEX	243-265	54
D1S439R	GTTGAAATGGTGAATTTG			
D1S1644F	GGATTGAGCCACAATACCAG	FAM	253-281	52-62
D1S1644R	ATACCTATGGAGGTGCACCA			
D1S1617F	AGAGCAAGACTCCATCGCAA	TET	118-133	52-54
D1S1617R	ACTTGTGGCTTCATGTTGGT			
D1S103F	ACGAACATTCTACAAGTTAC	FAM	82-90	-
D1S103R	TTTCAGAGAAACTGACCTGT			
D1S225F	GAGGCCTGGTAAAAAGTGAT	FAM	11-133	58-62
D1S225R	AGCAAGACTCCATCTCTAAATG			
D1S251F	GTCTCCAGCCTGCCAC	FAM	249-271	58-60
D1S251R	GACCAAGCAACTTCACTCC			
D1S2709F	TCATACCACATATCAGAATGTC	FAM	191-197	54
D1S2709R	ATCAATCAGTATCTAATAGCATCA			
D1S459F	AGCTGGGGTGTTTTATGGAG	HEX	177-193	56-60
D1S459R	GGGGACCCTGTTTCTGCTAC			
B804F	ATGGGCTCTGTCTCTATTGTTTC	FAM	143	50-60
B804R	CTATGTTCAGGCCTGCTATTTG			
B109F	CACCTCAGCCTCCCAAGTAGC	FAM	176	-
B109R	TCCAGCCTGAACAACATAGTGAAA			
B895F	TTTGTTTGAATCTCTGGTTATTTT	FAM	283	56-60
B895R	ACGTGGATGTGGTAGGATGAA			
D1S179F	GATCCTTTTATAGAAGATAGT	TET	163-193	52
D1S179R	TGTACGATGCATAGGTTTGC			
B781F	TATTCCTGGATTGTGTTTTTCGTC	FAM	230	60
B781R	AGGCCCAGGTAGTTTTAGATTCAC			
D1S235F	CAGCAAGAGTTCATGGGA	FAM	175-195	52-54

D1S235R	AACAGTCAATTACAAAATATGTGTG			
D1S2850F	CGAAGGTGTACTGGGACTGG	TET	145-153	58-62
D1S2850R	AATCAGGATCATGCTACAGGG			
B940F	TTTTCTGAGTGATTACCATAGTCC	FAM	307	54
B940R	ATAAGATACCAAATTCCAACCTG			
B519F	ATGGCTCTGGCAAATGT	HEX	276	50-56
B519R	TGGAGTGGGTTCAAGAGA			
B44F	CTCCAATTATTGCTCTGTCCTAC	HEX	191	58
B44R	GCACTATTTTGGGCACTATGG			
B541F	GTGCCGTGGGTACTAAAGACA	TET	282	50-60
B541R	AGACACTGCGAACTGAAATGG			
B193F	AGTTATCCACATTTCCCTCTTCTG	TET	279	52
B193R	GACTTTTATAGCTGCCAACACTGA			
B307F	TAATAGAGGGCTGAAAGGTGAAC	FAM	328	56-58
B307R	TAGGGCCAAAAGGAGGAA			
D1S517F	GTACATAGTGGAATGCCTGG	TET	213	60
D1S517R	GTAACTCACAAAGGCAAAGAC			
B359F	GCAGGATTGGCCTTTGATAA	HEX	235	58
B359R	CAAATTTGGGGGTGTAGCTC			
D1S1149F	GGAAAAATAGTGAGACCCCAT	TET	322	58
D1S1149R	AGTGGAGGTCCCTGAAGAC			
B177F	GGAAACCAGAAAATTCGTGTAA	TET	230	56-58
B177R	TTGCAACATCAACTCCTACCA			
D1S2785F	CGTGAATATCCTCAGGGAAT	FAM	148-178	52-56
D1S2785R	ATTGTGGCACCGTACTCC			
D1S180F	TCCCTAAAAGACTGCAAGCT	HEX	170	52-56
D1S180R	ACAGAGTCAAACTGTTGTGG			
D1S547F	CTGAAGTGGGAGGATTGCTT	FAM	282-308	56
D1S547R	AATTCAGGGGAGTTCCAGAG			
D1S2842F	TCACCTGACCTGTCCC	TET	217-231	48-54
D1S2842R	TGGTTCTCAGCCACAA			
D1S1609F	ATACATTTATATTTGCATCTGTGTG	FAM	180-208	48-60
D1S1609R	TGATTCTAGGTTGGTCACTCG			
D1S2682F	CCTTCTGCACCGTAACAC	FAM	117-145	48-52
D1S2682R	CCCCTTAAAGGAACTTGTC			
D1S3739F	GGAGTTAAGGTTGAAGAGCC	TET	155	52-58
D1S3739R	TTCACGTACAACAGTATCTC			

Primers used for DOP-PCR and amino-linking

obtained from Oswel, Southampton

Primer name	Primer sequence	Modification	PCR temp (°C)
DOP-primer 1	CCGACTCGAGNNNNNNCTAGAA		62
DOP-primer 2	CCGACTCGAGNNNNNNTAGGAG		62
DOP-primer 3	CCGACTCGAGNNNNNNTTCTAG		62
Amino-PCR	GGAAACAGCCCGACTCGAG	5'Amino-group	60

Primers used for RT-PCR of nidogen-1

obtained from Qiagen Operon, Crawley, UK

Primer name	Primer sequence	Amplicon size (bp)	PCR temp (°C)
NIDOGEN-1F	TGGTTGCATTCAGTCAAGGT	212	58
NIDOGEN-1R	CTCTGCCCCATCTTCAGTTC		

Primers used for dHPLC and sequencing analysis of nidogen-1

obtained from Qiagen Operon, Crawley, UK

Primer name	Primer sequence	Amplicon sixe (bp)	PCR temp (°C)	dHPLC temp (°C)
NIDEXON1F	CGCTCCCCTCCCCTATT	453	52-58	66/69
NIDEXON1R	GCCTGCTACGCCCAAGT			
NIDEXON2F	GCATGGGTACAGCTGACAGT	472	56-60	60/63
NIDEXON2R	ACAAGGCGTGAGACCAAAGT			
NIDEXON3F	CGCTTTATAGATAGTCTGGCGAAT	341	56-60	55/57/60
NIDEXON3R	GAATCTACTGAAGAACAAAGAAAGCA			
NIDEXON4AF	CTTGGGTAGTGGTGGCTGTT	299	56-60	62/63
NIDEXON4AR	GGCACACTGTATGTGTCAGCA			
NIDEXON4BF	TATGACCTGGCGACCACTC	300	56-60	59/62/64
NIDEXON4BR	GAAGCCCACACATCCAAGAT			
NIDEXON5F	GGGCTAGAGCTGCATGAAAT	235	56-60	59/60/64
NIDEXON5R	CCGCATGAAACGAACATAGA			
NIDEXON5AF	CCTGGTCCCTTGAATTGC	309		-
NIDEXON5AR	CCGTATAGCCAGCGACACA			
NIDEXON5BF	CCAGACGTGTGCTAACAACAG	283	56-60	-
NIDEXON5BR	CACCTACGGAGATGTGGGTAA			
NIDEXON6F	TGTTTGAAAGACTGGCTTGG	395	60	60/61/62
NIDEXON6R	TCAGAGTTCTCCTTCCATCCA			
NIDEXON7F	ACCAACCCCACTCCACACT	384	58-62	64/65
NIDEXON7R	CAGAGACAGCCCCAGTTCAC			
NIDEXON7AF	CCAGTTCTATGTTCTACAGCCAAA	360	56-60	-
NIDEXON7AR	AGGAGCCGAACGGAATCT			
NIDEXON7BF	AGGCTGAGGTGACCTTCGT	441	58-62	-
NIDEXON7BR	TCCCTCTCCAGATCCCAGA			
NIDEXON8F	AACCATCTTCCCTCTGTCTCC	380	50	62/63/64
NIDEXON8R	GAGACCAAAAACAATGTCAAAGA			
NIDEXON9F	CTGTGGAAGGAGGAAACCTG	244	60	61/62/64
NIDEXON9R	CAGCCCCAACAAGTCAGAAT			
NIDEXON9AF	CTGGTGTCCCATTCCACTCT	327	54-60	-
NIDEXON9AR	GCACTCGCAGGTGAACTGT			
NIDEXON9BF	CCTGCTACATCGGCACTCAT	348	52-60	-
NIDEXON9BR	CCTGGCATGTCCACCAGTAT			
NIDEXON10F	TTGATCTATTCTTAGGTGACAATGCT	323	58-62	54/60
NIDEXON10R	AACTTGACCAGACTCATTTTACCTG			

NIDEXON11F	GCTCAGGGGACTTTCTTCCT	284	50-60	60/61/63
NIDEXON11R	TGAGTTGGTGATTGGCACAT			
NIDEXON12F	GGTTCCCATAGATTTGAGTAAGTGG	281	58-62	56/60
NIDEXON12R	TGTGCTCTGCATTCATGGAC			
NIDEXON13F	ACGGATGACGCACCTGAC	369	56-60	63/66/67
NIDEXON13R	AGGTCCAGCAGAAGAGCAAA			
NIDEXON14F	GCATCTGTGTGTTCCTGGTG	295	50-60	62/63/64
NIDEXON14R	CCTGCTTGTTTTTCCAGGTC			
NIDEXON15F	AGGCACGCAAAAAGATTCTC	292	60	58/59
NIDEXON15R	TTTGGCTTCATTTGCTGTTC			
NIDEXON16F	CACCATGTGAAATAACATTGAGTAT	322	56-60	56/58/61
NIDEXON16R	CTCTAAAAGGGCAGATGCAA			
NIDEXON17F	AATCACTTCATTGTAATTCTTTTCAAG	354	56-60	55/60/61
NIDEXON17R	CAGGTGGCTGGTTAGATTTG			
NIDEXON18F	CCCACTCTTGCCCTCATTAC	281	56-60	58/63
NIDEXON18R	CTGGCCTCCCAGAGACAGT			
NIDEXON19F	AAAGAGGCTGCCTTGTTTAGA	245	50-62	59/61/62
NIDEXON19R	GCTCTGAGCCTCCATGTG			
NIDEXON20AF	GGCTATGGAGTCAGATCAGAGA	390	50-52	60/61/62
NIDEXON20AR	CATGTTTTGGGGAACAGTGA			

Antibodies used in this thesis

Primary antibodies

Antibody	Source	Dilution	Antigen unmasking	Raised in
Nidogen-1	Calbiochem	1µl in 300µl PBS	3 minutes trypsinisation	Rabbit
Laminin-1	Novo castra	1µl in 100µl PBS	3 minutes trypsinisation	Mouse
Collagen type IV	Novo castra	1µl in 100µl PBS	3 minutes trypinisation + 10 minutes boiling sodium citrate	Mouse

Secondary antibody and steptavidin detection

Antibody	Source	Dilution
Anti-rabbit	Vector laboratories	1µl in 800µl PBS
Anti-mouse	Vector laboratories	1µl in 100µl PBS
Steptavadin horseradish peroxidase	Vector laboratories	1µl in 200µl PBS

Appendix ii

Genes mapping to the 1q42-44 region

(Taken from http://www.ncbi.nlm.nih.gov/; 10/2004)

start	stop	Cyto	Description
227084993	227136463	1q42.2	component of oligomeric golgi complex 2
227145020	227156602	1q42-q43	angiotensinogen (serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 8)
227189865	227244243	1q42.11-q42.3	calpain 9 (nCL-4)
227279601	227311012	1q42.13-q43	hypothetical protein FLJ14525
227310992	227317273	1q42.2	LOC440729
227348725	227421319	1q42.2	tetratricopeptide repeat domain 13
227421558	227443214	1q42.2	likely ortholog of yeast ARV1
227461670	227482720	1q42.2	hypothetical gene supported by BC009447
227483511	227605268	1q42.2	similar to peptidylprolyl isomerase A isoform 1; cyclophilin A; peptidyl-prolyl cis-trans isomerase A;
227605409	227657921	1q42.2	similar to TRIM9-like protein TNL
227666247	227683651	1q42.2	hypothetical protein DKFZp547B1713
227683711	227720268	1q42.11-42.3	glyceronephosphate O-acyltransferase
227775217	227780288	1q42.2	exocyst complex component 8
227780466	227796724	1q42.12-q43	hypothetical protein DKFZp547N043
227808166	227867525	1q42.1	egl nine homolog 1 (C. elegans)
227969814	227970953	1q42.2	LOC440731
227971134	228009005	1q42.1	translin-associated factor X
228069296	228483608	1q42.1	disrupted in schizophrenia 1
228840449	228957171	1q42.2	signal-induced proliferation-associated 1 like 2
229247505	229250886	1q42.2	KIAA1383 protein
229393105	229420954	1q42.2	hypothetical protein MGC13186
229426617	229701827	1q42.2	hypothetical protein FLJ11383
229731286	229737959	1q42.2	similar to 40S ribosomal protein S7 (S8)
229770249	229826486	1q42	mixed lineage kinase 4
230056485	230114719	1q42-q43	potassium channel, subfamily K, member 1
230347414	230766997	1q42.2	solute carrier family 35, member F3
230799137	230801001	1q42.2	similar to ribosomal protein S15; rat insulinoma gene
230815948	230826299	1q42.2	chromosome 1 open reading frame 31
230833794	230921584	1q42.3	TAR (HIV) RNA binding protein 1
231049513	231052006	1q42.3	interferon regulatory factor 2 binding protein 2
231399828	231406481	1q42.3	hypothetical LOC284527
231476438	231485672	1q42.3	similar to Translocase of outer mitochondrial membrane 20 homolog

231488266	231517871	1q42.3	KIAA0117 protein
231598701	231618172	1q42	translocase of outer mitochondrial membrane 20 homolog (yeast)
231656253	231817541	1q42.1-q43	AT rich interactive domain 4B (RBP1- like)
231817926	231832451	1q43	geranylgeranyl diphosphate synthase 1
231856812	231938321	1q42.3	tubulin-specific chaperone e
231939288	231993822	1q42.3	beta 1,3-N-acetylgalactosaminyltransferase-II
232040174	232140095	1q42.3	guanine nucleotide binding protein (G protein), gamma 4
232150384	232323008	1q42.1-q42.2	Chediak-Higashi syndrome 1
232466152	232554503	1q43	**nidogen (enactin)**
232599402	232601713	1q42.3	similar to aconitase 2 precursor; aconitate hydratase; citrate hydro-lyase
232631915	232698247	1q42-q43	transmembrane 7 superfamily member 1 (upregulated in kidney)
232706502	232771326	1q42.2-q43	ERO1-like beta (S. cerevisiae)
232883721	232974049	1q42.3-q43	EDAR-associated death domain
232972423	232974247	1q43	enolase 1, (alpha) pseudogene
233007606	233039037	1q42-q43	lectin, galactoside-binding, soluble, 8 (galectin 8)
233038348	233083438	1q43	protein BAP28
233085113	233087322	1q43	similar to Protein BAP28
233089913	233093845	1q43	similar to Protein BAP28

Start	Stop	Cyto	Description
233175840	233253281	1q42-q43	actinin, alpha 2
233284759	233390003	1q43	5-methyltetrahydrofolate-homocysteine methyltransferase
233470651	233495386	1q43	similar to 60S ribosomal protein L35
233531743	234322694	1q42.1-q43	ryanodine receptor 2 (cardiac)
234371751	234379976	1q43	zona pellucida glycoprotein 4
234416091	234416896	1q43	similar to actin, gamma, cytoplasmic
234969725	234975358	1q43	hypothetical protein LOC339535
234979602	235384209	1q43	similar to Keratin, type I cytoskeletal 18 (Cytokeratin 18) (K18) (CK 18)
236396793	236398565	1q41-q44	cholinergic receptor, muscarinic 3
236460221	236463588	1q43	similar to Hydroxymethylglutaryl-CoA synthase,
236502010	236964518	1q43	formin 2
236631871	236634716	1q43	proteasome 26S non-ATPase subunit 2 pseudogene

236979831	237101430	1q43	gremlin 2 homolog, cysteine knot superfamily (Xenopus laevis)
237264912	237846519	1q43	regulator of G-protein signalling 7
237900021	237923588	1q43	hypothetical LOC388755
237986947	238009095	1q42.1	fumarate hydratase
238020286	238084984	1q42-q44	kynurenine 3-monooxygenase (kynurenine 3-hydroxylase)
238082956	238129652	1q43	opsin 3 (encephalopsin, panopsin)
238122971	238125273	1q42-qter	choroideremia-like (Rab escort protein 2)
238172923	238290810	1q43	hypothetical protein FLJ32978
238313230	238322130	1q43	similar to 60S ribosomal protein L6 (TAX-responsive enhancer element binding protein 107)
238337576	238379089	1q42-q43	exonuclease 1
238447110	238483026	1q43	similar to Bcl-2-interacting protein beclin
238484833	238488416	1q43	similar to MAP1 light chain 3-like protein 2
238546428	238549618	1q43	HSA1q43-44 beta-tubulin 4Q (TUBB4Q) pseudogene
238578313	239014039	1q43	hypothetical protein FLJ40773
238866265	238866804	1q43	similar to zinc finger protein 532
239093925	239262487	1q43	similar to ribosomal protein L24-like; 60S ribosomal protein L30 isolog; my024 protein;
239537075	239542789	1q44	similar to bA476I15.3 (novel protein similar to septin)
239542926	239546379	1q44	similar to hypothetical protein LOC349114
239546427	239556551	1q44	similar to hypothetical protein LOC284701
239577187	239591067	1q44	hypothetical LOC339077
239613772	239714664	1q44	KARP-1-binding protein
239745517	239978483	1q44	serologically defined colon cancer antigen 8
239977576	240332594	1q43-q44	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)
240406745	240536660	1q44	hypothetical protein LOC339529
240540626	240546819	1q44-qter	zinc finger protein 238
240553673	240557252	1q44	hypothetical gene supported by AK124614
240596708	240599386	1q44	hypothetical LOC343484
240693879	240878431	1q44	hypothetical LOC200159
240884318	240898072	1q44	**similar to TGFB-induced factor 2; TGF(beta)-induced transcription factor 2;**
240950731	241129703	1q44	hypothetical protein MGC33370
241087733	241088030	1q44	cytochrome c, somatic pseudogene
241142415	241198376	1q44	CGI-146 protein
241325004	241332999	1q44	family with sequence similarity 36, member A

241343004	241353868	1q44	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)
241459733	241577179	1q44	hypothetical protein MGC12458
241843021	241861060	1q44	hypothetical protein MGC35030
242000345	242192597	1q44	hypothetical protein FLJ10157
242238685	242844440	1q44	SET and MYND domain containing 3
242998178	243007704	1q44	LOC441928
243008283	243011266	1q44	similar to XRnf12C
243029909	243055604	1q44	transcription factor B2, mitochondrial
243055847	243156467	1q44	hypothetical protein FLJ32001
243213766	243257480	1q44	CGI-49 protein
243278437	243281732	1q44	hypothetical protein LOC149134
243296560	243319402	1q44	similar to kinesin-like protein (103.5 kD) (klp-6)
243328714	243420767	1q44	ELYS transcription factor-like protein TMBS62
243434905	243489418	1q44	zinc finger protein SBZF3
243526128	243528880	1q44	hypothetical protein MGC12466
243589334	243591461	1q44	hypothetical protein FLJ12606
243593854	243598904	1q44	hypothetical LOC400812
243599503	243601760	1q44	FLJ45717 protein
243645244	243649156	1q44	zinc finger protein 124 (HZF-16)
243673290	243691438	1q44	similar to Zinc finger protein 492
243719283	243721240	1q44	similar to Heat shock cognate 71 kDa protein
243726568	243727194	1q44	similar to putative G-protein coupled receptor
243745415	243746488	1q44	vomeronasal 1 receptor 5
243789664	243821086	1q44	zinc finger protein 496
243818923	243822118	1q44	hypothetical LOC400813
243907392	243938451	1q44	cold autoinflammatory syndrome 1
243940372	243941325	1q44	olfactory receptor, family 2, subfamily B, member 11
243960649	244007806	1q44	similar to olfactory receptor Olr1657; similar to olfactory receptor MOR256-12
244019475	244023182	1q44	olfactory receptor, family 2, subfamily C, member 3
244038490	244065900	1q44	LOC148823
244077709	244109905	1q44	olfactory receptor, family 2, subfamily G, member 3
244109024	244109902	1q44	olfactory receptor, family 5, subfamily AV, member 1 pseudogene
244161461	244162384	1q44	similar to Olfactory receptor 13G1

244212242	244213586	1q44	olfactory receptor, family 5, subfamily AX, member 1
244227958	244228902	1q44	olfactory receptor, family 5, subfamily AY, member 1
244246805	244248418	1q44	olfactory receptor, family 1, subfamily C, member 1
244264027	244273033	1q44	VNFT9373
244264253	244265176	1q44	olfactory receptor, family 9, subfamily H, member 1 pseudogene
244304143	244331239	1q44	olfactory receptor, family 11, subfamily L, member 1
244322660	244323588	1q44	olfactory receptor, family 6, subfamily R, member 1 pseudogene
244346542	244366917	1q44	BIA2
244384930	244385874	1q44	olfactory receptor, family 2, subfamily W, member 3
244410361	244411299	1q44	similar to seven transmembrane helix receptor
244423220	244424162	1q44	olfactory receptor, family 2, subfamily AJ, member 1
244426534	244590265	1q44	hypothetical protein MGC40047
244428414	244429158	1q44	olfactory receptor, family 2, subfamily X, member 1 pseudogene
244438201	244455682	1q44	olfactory receptor, family 2, subfamily AK, member 2
244464085	244465018	1q44	olfactory receptor, family 2, subfamily L, member 9 pseudogene
244464085	244480549	1q44	olfactory receptor, family 2, subfamily L, member 1 pseudogene
244492474	244493409	1q44	olfactory receptor, family 2, subfamily L, member 6 pseudogene
244511291	244512226	1q44	olfactory receptor, family 2, subfamily L, member 5
244527611	244528549	1q44	olfactory receptor, family 2, subfamily L, member 2
244550025	244550963	1q44	olfactory receptor, family 2, subfamily L, member 3
244572979	244573939	1q44	olfactory receptor, family 2, subfamily T, member 32 pseudogene
244611479	244612123	1q44	olfactory receptor, family 2, subfamily M, member 1 pseudogene
244634491	244635429	1q44	olfactory receptor, family 2, subfamily M, member 5

Appendix iii

Genomic and amino acid sequence of nidogen-1

Genomic sequence of nidogen-1

Taken from http://www.ensembl.org/. Sequence changes identified in section 5.3.2 are highlighted in red

Exon number	Exon / Intron	Length	Sequence
	5' upstream sequence		.....acatccccgccttcctctgtcctggccgcgggaccgggtttgcgggaccg
1	ENSE00001342116	237 bp	CAGTTCGGGAACATGTTGGCCTCGAGCAGCCGGATCCGGGCTGCGTGGACGCGGGCGCTG CTGCTGCCGCTGCTGCTGGCGGGGCCTGTGGGCTGCCTGAGCCGCCAGGAGCTCTTTCCC TTCGGCCCCGGACAGGGGGACCTGGAGCTGGAGGACGGGGATGACTTCGTCTCTCCTGCC CTGGAGCTGAGTGGGGCGCTCCGCTTCTACGACAGATCCGACATCGACGCAGTCTAC
	Intron 1-2	15865 bp	gtgagtgagccccgggagggcgggc.....tttaaccacacttggttttgtacag
2	ENSE00001022663	300 bp	GTCACCACAAATGGCATCATTGCTACGAGTGAACCCCCGGCCAAAGAATCCCATCCCGGG CTCTTCCCACCAACATTCGGTGCAGTCGCCCCTTTCCTGGCGGACTTGGACACGACCGAT GGCCTGGGGAAGGTTTATTATCGAGAAGACTTATCCCCCTCCATCACTCAGCGAGCAGCA GAGTGTGTCCACAGAGGGTTCCCGGAGATCTCTTTCCAGCCTAGTAGCGCGGTGGTTGTC ACTTGGGAATCCGTGGCCCCCTACCAAGGGCCCAGCAGGGACCCAGACCAGAAAGGCAAG
	Intron 2-3	3006 bp	gtaagctcccctccaggtccaagtg.....actgaattttaactcatcttcacag
3	ENSE00001022662	227 bp	AGAAACACGTTCCAGGCTGTTCTAGCCTCCTCTGATTCCAGCTCCTATGCCATTTTCCTT TATCCTGAGGATGGTCTGCAGTTCCATACGACATTCTCAAAGAAGGAAAACAACCAAGTT CCTGCCGTGGTTGCATTCAGTCAAGGTTCAGTGGGATTCTTATGGAAGAGCAACGGAGCT TATAACATATTTGCTAATGACAGGGAATCAGTTGAAAATTTGGCCAA
	Intron 3-4	3164 bp	gtatgctttctttgttcttcagtag.....ctaagaagcctgtccttcctcctag
4	ENSE00001022668	383 bp	GAGTAGTAACTCTGGGCAGCAGGGTGTCTGGGTGTTTGAGATTGGGAGTCCAGCCACCAC CAATGGCGTGGTGCCTGCAGACGTGATCCTCGGAACTGAAGATGGGGCAGAGTATGATGA TGAGGATGAAGATTATGACCTGGCGACCACTCGTCTGGGCCTGGAGGATGTGGGCACCAC GCCCTTCTCCTACAAGGCTCTGAGAAGGGGAGGTGCTGACACATACAGTGTGCCCAGCGT CCTCTCCCCGCGCCGGGCAGCTACCGAAAGGCCCCTTGGACCTCCCACAGAGAGAACCAG GTCTTTCCAGTTGGCAGTGGAGACTTTTCACCAGCAGCACCCTCAGGTCATAGATGTGGA TGAAGTTGAGGAAACAGGAGTTG
	Intron 4-5	3656 bp	gtaagaccatttaagttgcagttgg.....gtgtgcaattgttttcttggaccag
5	ENSE00001022661	150 bp	TTTTCAGCTATAACACGGATTCCCGCCAGACGTGTGCTAACAACAGACACCAGTGCTCGG

			TGCACGCAGAGTGCAGGGACTACGCCACGGGCTTCTGCTGCAGCTGTGTCGCTGGCTATA
			CGGGCAATGGCAGGCAATGTGTTGCAGAAG
	Intron 5-6	5451 bp	gtaatttgcttttctatgttcgttt.....attctttctttgttttcttgctcag
6	ENSE00001022658	252 bp	GTTCCCCCCAGCGAGTCAATGGCAAGGTGAAAGGAAGGATCTTTGTGGGGAGCAGCCAGG
			TCCCCATTGTCTTTGAGAACACTGACCTCCACTCTTACGTAGTAATGAACCACGGGCGCT
			CCTACACAGCCATCAGCACCATTCCCGAGACCGTTGGATATTCTCTGCTTCCACTGGCCC
			CAGTTGGAGGCATCATTGGATGGATGTTTGCAGTGGAGCAGGACGGATTCAAGAATGGGT
			TCAGCATCACCG
	Intron 6-7	2650 bp	gtaatttatatccgacattaaaatt.....gtgacagtctcatcttgtttttcag
7	ENSE00001022659	201 bp	GGGGTGAGTTCACTCGCCAGGCTGAGGTGACCTTCGTGGGGCACCCGGGCAATCTGGTCA
			TTAAGCAGCGGTTCAGCGGCATCGATGAGCATGGGCACCTGACCATCGACACGGAGCTGG
			AGGGCCGCGTGCCGCAGATTCCGTTCGGCTCCTCCGTGCACATTGAGCCCTACACGGAGC
			TGTACCACTACTCCACCTCAG
	Intron 7-8	3408 bp	gtgagccccagctgtgtcccagggc.....cccctccatccttcccactctgcag
8	ENSE00000793467	246 bp	TGATCACTTCCTCCTCCACCCGGGAGTACACGGTGACTGAGCCCGAGCGAGATGGGGCAT
			CTCCTTCACGCATCTACACTTACCAGTGGCGCCAGACCATCACCTTCCAGGAATGCGTCC
			ACGATGACTCCCGGCCAGCCCTGCCCAGCACCCAGCAGCTCTCGGTGGACAGCGTGTTCG
			TCCTGTACAACCAGGAGGAGAAGATCTTGCGCTATGCTCTCAGCAACTCCATTGGGCCTG
			TGAGGG
	Intron 8-9	1682 bp	gtaaggggatacagctcatgctggg.....ggttccaatgtctctgtcttcacag
9	ENSE00000793466	144 bp	AAGGCTCCCCTGATGCTCTTCAGAATCCCTGCTACATCGGCACTCATGGGTGTGACACCA
			ACGCGGCCTGTCGCCCTGGTCCCAGGACACAGTTCACCTGCGAGTGCTCCATCGGCTTCC
			GAGGAGACGGGCGAACCTGCTATG
	Intron 9-10	6796 bp	gtacagcctttgattctgacttgtt.....cagtaaaaaaaaattttttgctgcag
10	ENSE00000793465	126 bp	ATATTGATGAATGTTCAGAACAACCCTCAGTGTGTGGGAGCCACACAATCTGCAATAATC
			ACCCAGGAACCTTCCGCTGCGAGTGTGTGGAGGGCTACCAGTTTTCAGATGAGGGAACGT
			GTGTGG
	Intron 10-11	3587 bp	gtaagttctccaggtaacttttcga.....tgtgcatcaacctttttgttctgaag
11	ENSE00000793464	150 bp	CTGTCGTGGACCAGCGCCCCATCAACTACTGTGAAACTGGCCTTCATAACTGCGACATAC
			CCCAGCGGGCCCAGTGTATCTACACAGGAGGCTCCTCCTACACCTGTTCCTGCTTGCCAG
			GCTTTTCTGGGGATGGCCAAGCCTGCCAAG
	Intron 11-12	1367 bp	gtgtgtggttgctctgatcttcccc.....cctgggcctggtggctgttttacag
12	ENSE00000793463	123 bp	ATGTAGATGAATGCCAGCCAAGCCGATGTCACCCTGACGCCTTCTGCTACAACACTCCAG
			GCTCTTTCACGTGCCAGTGCAAACCTGGTTATCAGGGAGACGGCTTCCGTTGCGTGCCCG
			GAG
	Intron 12-13	18048 bp	gtaaggtggtgggacatctggagtc.....ttcttgttgcctgtttctctgatag
13	ENSE00000793462	228 bp	AGGTGGAGAAAACCCGGTGCCAGCACGAGCGAGAACACATTCTCGGGGCAGCGGGGGCGA
			CAGACCCACAGCGACCCATTCCTCCGGGGCTGTTCGTTCCTGAGTGCGATGCGCACGGGC

			ACTACGCGCCCACCCAGTGCCACGGCAGCACCGGCTACTGCTGGTGCGTGGATCGCGACG
			GCCGCGAGGTGGAGGGCACCAGGACCAGGCCCGGGATGACGCCCCCGT
	Intron 13-14	2586 bp	gtaagtgggtgcaggccgtgcttgg.....ctcactgacccccctctgccatgcag
14	ENSE00000793461	173 bp	GTCTGAGTACAGTGGCTCCCCCGATTCACCAAGGACCTGCGGTGCCTACCGCCGTGATCC
			CCTTGCCTCCTGGGACCCATTTACTCTTTGCCCAGACTGGGAAGATTGAGCGCCTGCCCC
			TGGAGGGAAATACCATGAGGAAGACAGAAGCAAAGGCGTTCCTTCATGTCCCG
	Intron 14-15	5380 bp	gtgagtgctggctgattatacctgg.....tctaaccactggtcttcacatccag
15	ENSE00000793460	127 bp	GCTAAAGTCATCATTGGACTGGCCTTTGACTGCGTGGACAAGATGGTTTACTGGACGGAC
			ATCACTGAGCCTTCCATTGGGAGAGCTAGTCTACATGGTGGAGAGCCAACCACCATCATT
			AGACAAG
	Intron 15-16	3596 bp	gtaagtaagcaaatatcatttttcc.....agaggagctctgtgtttacttctag
16	ENSE00000793459	172 bp	ATCTTGGAAGTCCAGAAGGTATCGCTGTTGATCACCTTGGCCGCAACATCTTCTGGACAG
			ACTCTAACCTGGATCGAATAGAAGTGGCGAAGCTGGACGGCACGCAGCGCCGGGTGCTCT
			TTGAGACTGACTTGGTGAATCCCAGAGGCATTGTAACGGATTCCGTGAGAGG
	Intron 16-17	957 bp	gtatctttgtgtaaatatgtgtgtg.....ctttccccccttttttcctcctccag
17	ENSE00000793458	158 bp	GAACCTTTACTGGACAGACTGGAACAGAGATAACCCCAAGATTGAAACTTCCTACATGGA
			CGGCACGAACCGGAGGATCCTTGTGCAGGATGACCTGGGCTTGCCCAATGGACTGACCTT
			CGATGCGTTCTCATCTCAGCTCTGCTGGGTGGATGCAG
	Intron 17-18	550 bp	gtgatggaaagctggacactgtcca.....aatgaattgtttccactccacacag
18	ENSE00000793457	124 bp	GCACCAATCGGGCGGAATGCCTGAACCCCAGTCAGCCCAGCAGACGCAAGGCTCTCGAAG
			GGCTCCAGTATCCTTTTGCTGTGACGAGCTACGGGAAGAATCTGTATTTCACAGACTGGA
			AGAT
	Intron 18-19	714 bp	gtacgtcagctgtagccatgggggc.....ccttctgccctttggtgctttgtag
19	ENSE00000793456	113 bp	GAATTCCGTGGTTGCTCTCGATCTTGCAATTTCCAAGGAGACGGATGCTTTCCAACCCCA
			CAAGCAGACCCGGCTGTATGGCATCACCACGGCCCTGTCTCAGTGTCCGCAAG
	Intron 19-20	1006 bp	gtaactctcctgtgctgttactgtt.....ttgaactgatttctgccttttccag
20	ENSE00001342076	2148 bp	GCCATAACTACTGCTCAGTGAACAATGGCGGCTGCACCCACCTATGCTTGGCCACCCCAG
			GGAGCAGGACCTGCCGTTGCCCTGACAACACCTTGGGAGTTGACTGTATCGAACAGAAAT
			GAAGACAAGAGTGCCTTATTTCCTTTCCAAGTATTTCACAGCAACACTCTACTTGAAGCA
			ACTTGGTCCAGATTGAAAAGTGTCCTCTGGCTGAGTGGCCACTAGGCCCAGACCCAGCCC
			AGCCTGAGCCCCAACAACTTTTCCCTCACTGTTCCCCAAAACATGCACCCTGGACTTCTC
			TAATAGAAAAGTCTCCACCCCTACACAAGGACAGAACCCTCCACCCCTACCCCCAACCCT
			CAGACAGACTTATACACCCCTGAGTGAGGATTACATGCCCATCCCAGTGTCCTAGGACCT
			TTTCCCAATACTAGCCCCCCAGTGGTGAACAGAACCTCCCAAATTTGAGTTGCACCCTTC
			CCTGTGGCCTTATGAGCTCAGCCTCGCTTTGAGGTACCCACCGTCCTGTCAGCTCCTTGA
			CCTATGAGCCGGGGCCTGACTAGGAAAAGTTGGGAGTTAAGGAGGAAATTAGCATTCCTT
			AATGTTTTGTTTTGGTGCTCTGAATTTCTTCTTTATTATAGTCCTATAGTTTTACTCCTC
			AGTTCCTCACCATCATCATCTTGTCTAAGACCCCCATTATAATATTCATGCGCTGCTTTT


TCATCAAAACCTACCCTGTCCTAGAGATCTATGGGCATTTGGTGGATGATAATGAGCAGC
CCCTCCCAGATAGAATGTCAATATTTGAGCAGTAGGATATTGGCATTTGTTAGTTAAAGG
CTTAAATCAAAAGAATGTCCAATGGTAGGAATTTCAAGGTGTAGGTCAGATATTTGAGAA
TAGGGGATTTTTTTGATGTGCCTTAAATTATACCAAAGATTACTAATTATTCCTCTTTGC
CCAAAATACTTGCATCCAAGGTTCTAGTCTCTGTTGCTGTGCTGGTCTTTAGCCCCACTG
CTTGCACTGATGTCCCTCCTTTTCACGGAGACCTATCTGAGGTACAGGATGGGGCTGGCA
CCAGATGATGTCCCACCACAGTCCCTCACCTCCGGCCTCCACATGACAGAACCAATTTAC
ACTCAACCATGACCTCACCCCTCCTTGGTTTCTCCCTCGATCTGTGGCCCTTTTTGGATG
TATTCTTATCTAACAACACAATCCGGAAAGACTGAATTGAATATTTATACTAATGGTTCA
TATCCTTTATTGCTCAATGATCTAATTAAAGGGATCATTGCCACATTTCATGTTTATATT
TCTACAATTTGTTTAGAAAACATCTCCTGACCATATCAGTAGCTCGTGTTATCTTTTTAT
CAACTGCTTCCCAGAGTCCTAAAACAATAGAAATTTTGGATTGAAAAGTTCAGCATAAGG
AGTTTGAGTCAGTAAAGGATGGGATAAAGGAGTCGAGATGATTCAATGAAAAGTATCACA
AAAAAGAGATTGATCAACAAGAGAAATAAAAAAGCCCAAGAGGAAGTGGTAGGGGAAGGA
ATTTAAGAACAGCAATAAGTAAAACTCTTAAGTAACTCCAAAAAGAAAATGGTACATTTT
GCCAAAGACCACTTATACTTGAGAACATGGAAGAATTTGCCTGATACTCTCTTTGGGGAA
AAGAGTCTCTCCTCTTTTCCTCAAACCCCAGTACACTCAGCCTCTCTGCCCCACCTTCTC
CTGACTTTGTCCTCACTTGCTTCTGCAGTACATTGGAACCTGAATTGAAAGAAAGTCTTC
CTTGAATAATTGGAGTTTGTCTTGAGAGGCAAATATAGCCCCAAGAATCACAAGATTCGA
GGACCATGTAGGTCTTTTACGTAGCCCAAATCCATAAATTAGTCTCACTTTTTGTATTTA
TCGTTTCATATTAAACCCTCTATATCAAATGTTCATCATGATTTTGTATGATTTTTATAA
CTATTTTATTCATTTTATTAGATTTATTCTAAAATTTTTTAATGGTAAATTCTTAAACTG
TGGAAACCACTGAAGGTGCTTATTAACTGTTCTCCCAGATTTGTACAAGTATTGGATGAT
TCCTTGAGTTTACAGCTGTACAAATAGTGTGGAAAATAAACTTTTTTT
5' downstream sequence
aaaaaagaaaatcttcctctatcttttctgaatgaagagatttaatgttg

Amino acid sequence of nidogen-1

(taken from http://www.ensembl.org/)

GALRFYDRSDIDAVYVTTNGIIATSEPPAKESHPGLFPPTFGAVAPFLADLDTTDGLGKV
YYREDLSPSITQRAAECVHRGFPEISFQPSSAVVVTWESVAPYQGPSRDPDQKGKRNTFQ
AVLASSDSSSYAIFLYPEDGLQFHTTFSKKENNQVPAVVAFSQGSVGFLWKSNGAYNIFA
NDRESVENLAKSSNSGQQGVWVFEIGSPATTNGVVPADVILGTEDGAEYDDEDEDYDLAT
TRLGLEDVGTTPFSYKALRRGGADTYSVPSVLSPRRAATERPLGPPTERTRSFQLAVETF
HQQHPQVIDVDEVEETGVVFSYNTDSRQTCANNRHQCSVHAECRDYATGFCCSCVAGYTG
NGRQCVAEGSPQRVNGKVKGRIFVGSSQVPIVFENTDLHSYVVMNHGRSYTAISTIPETV
GYSLLPLAPVGGIIGWMFAVEQDGFKNGFSITGGEFTRQAEVTFVGHPGNLVIKQRFSGI
DEHGHLTIDTELEGRVPQIPFGSSVHIEPYTELYHYSTSVITSSSTREYTVTEPERDGAS
PSRIYTYQWRQTITFQECVHDDSRPALPSTQQLSVDSVFVLYNQEEKILRYALSNSIGPV
REGSPDALQNPCYIGTHGCDTNAACRPGPRTQFTCECSIGFRGDGRTCYDIDECSEQPSV
CGSHTICNNHPGTFRCECVEGYQFSDEGTCVAVVDQRPINYCETGLHNCDIPQRAQCIYT
GGSSYTCSCLPGFSGDGQACQDVDECQPSRCHPDAFCYNTPGSFTCQCKPGYQGDGFRCV
PGEVEKTRCQHEREHILGAAGATDPQRPIPPGLFVPECDAHGHYAPTQCHGSTGYCWCVD
RDGREVEGTRTRPGMTPPCLSTVAPPIHQGPAVPTAVIPLPPGTHLLFAQTGKIERLPLE
GNTMRKTEAKAFLHVPAKVIIGLAFDCVDKMVYWTDITEPSIGRASLHGGEPTTIIRQDL
GSPEGIAVDHLGRNIFWTDSNLDRIEVAKLDGTQRRVLFETDLVNPRGIVTDSVRGNLYW
TDWNRDNPKIETSYMDGTNRRILVQDDLGLPNGLTFDAFSSQLCWVDAGTNRAECLNPSQ
PSRRKALEGLQYPFAVTSYGKNLYFTDWKMNSVVALDLAISKETDAFQPHKQTRLYGITT
ALSQCPQGHNYCSVNNGGCTHLCLATPGSRTCRCPDNTLGVDCIEQK

Bibliography

Al-Awadi, S. A., Farag, T. I., Usha, R., el-Khalifa, M. Y., Sundareshan, T. S. and Al-Othman, S. A. (1986). Interstitial deletion of the long arm of chromosome 1 [del(1)(q32q42)]. *Am J Med Genet* **23**, 931-3.

Alexander, J., Rothenberg, M., Henry, G. L. and Stainier, D. Y. (1999). casanova plays an early and essential role in endoderm formation in zebrafish. *Dev Biol* **215**, 343-57.

Arrechedera, H., Alvarez, M., Strauss, M. and Ayesta, C. (1987). Origin of mesenchymal tissue in the septum primum: a structural and ultrastructural study. *J Mol Cell Cardiol* **19**, 641-51.

Aumailley, M., Battaglia, C., Mayer, U., Reinhardt, D., Nischt, R., Timpl, R. and Fox, J. W. (1993). Nidogen mediates the formation of ternary complexes of basement membrane components. *Kidney Int* **43**, 7-12.

Aumailley, M., Wiedemann, H., Mann, K. and Timpl, R. (1989). Binding of nidogen and the laminin-nidogen complex to basement membrane collagen type IV. *Eur J Biochem* **184**, 241-8.

Baldwin, H. S., Lloyd, T. R. and Solursh, M. (1994). Hyaluronate degradation affects ventricular function of the early postlooped embryonic rat heart in situ. *Circ Res* **74**, 244-52.

Barron, M., Gao, M. and Lough, J. (2000). Requirement for BMP and FGF signaling during cardiogenic induction in non-precardiac mesoderm is specific, transient, and cooperative. *Dev Dyn* **218**, 383-93.

Bartram, U., Molin, D. G., Wisse, L. J., Mohamad, A., Sanford, L. P., Doetschman, T., Speer, C. P., Poelmann, R. E. and Gittenberger-de Groot, A. C. (2001). Double-outlet right ventricle and overriding tricuspid valve reflect disturbances of looping, myocardialization, endocardial cushion differentiation, and apoptosis in TGF-beta(2)-knockout mice. *Circulation* **103**, 2745-52.

Benson, D. W., Silberbach, G. M., Kavanaugh-McHugh, A., Cottrill, C., Zhang, Y., Riggs, S., Smalls, O., Johnson, M. C., Watson, M. S., Seidman, J. G. et al. (1999). Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest* **104**, 1567-73.

Bonnet, D., Pelet, A., Legeai-Mallet, L., Sidi, D., Mathieu, M., Parent, P., Plauchu, H., Serville, F., Schinzel, A., Weissenbach, J. et al. (1994). A gene for Holt-Oram syndrome maps to the distal long arm of chromosome 12. *Nat Genet* **6**, 405-8.

Bradley, S. M. (1999). Neonatal repair and dealing with a single ventricle. *J S C Med Assoc* **95**, 335-8.

Bressan, G. M., Daga-Gordini, D., Colombatti, A., Castellani, I., Marigo, V. and Volpin, D. (1993). Emilin, a component of elastic fibers preferentially located at the elastin-microfibrils interface. *J Cell Biol* **121**, 201-12.

Brewer, C., Holloway, S., Zawalnyski, P., Schinzel, A. and FitzPatrick, D. (1998). A chromosomal deletion map of human malformations. *Am J Hum Genet* **63**, 1153-9.

Bruder, C. E., Hirvela, C., Tapia-Paez, I., Fransson, I., Segraves, R., Hamilton, G., Zhang, X. X., Evans, D. G., Wallace, A. J., Baser, M. E. et al. (2001). High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH. *Hum Mol Genet* **10**, 271-82.

Bruder, C. E., Ichimura, K., Blennow, E., Ikeuchi, T., Yamaguchi, T., Yuasa, Y., Collins, V. P. and Dumanski, J. P. (1999). Severe phenotype of neurofibromatosis type 2 in a patient with a 7.4-MB constitutional deletion on chromosome 22: possible localization of a neurofibromatosis type 2 modifier gene? *Genes Chromosomes Cancer* **25**, 184-90.

Buckley, P. G., Mantripragada, K. K., Benetkiewicz, M., Tapia-Paez, I., Diaz De Stahl, T., Rosenquist, M., Ali, H., Jarbo, C., De Bustos, C., Hirvela, C. et al.

(2002). A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. *Hum Mol Genet* **11**, 3221-9.

Burn, J., Brennan, P., Little, J., Holloway, S., Coffey, R., Somerville, J., Dennis, N. R., Allan, L., Arnold, R., Deanfield, J. E. et al. (1998). Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. *Lancet* **351**, 311-6.

Capdevila, J., Vogan, K. J., Tabin, C. J. and Izpisua Belmonte, J. C. (2000). Mechanisms of left-right determination in vertebrates. *Cell* **101**, 9-21.

Cartegni, L., Chew, S. L. and Krainer, A. R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **3**, 285-98.

Carter, N. P., Fiegler, H. and Piper, J. (2002). Comparative analysis of comparative genomic hybridization microarray technologies: report of a workshop sponsored by the Wellcome Trust. *Cytometry* **49**, 43-8.

Chan-Thomas, P. S., Thompson, R. P., Robert, B., Yacoub, M. H. and Barton, P. J. (1993). Expression of homeobox genes Msx-1 (Hox-7) and Msx-2 (Hox-8) during cardiac development in the chick. *Dev Dyn* **197**, 203-16.

Chen, J. N. and Fishman, M. C. (2000). Genetics of heart development. *Trends Genet* **16**, 383-8.

Christoffels, V. M., Habets, P. E., Franco, D., Campione, M., de Jong, F., Lamers, W. H., Bao, Z. Z., Palmer, S., Biben, C., Harvey, R. P. et al. (2000). Chamber formation and morphogenesis in the developing mammalian heart. *Dev Biol* **223**, 266-78.

Clark, B. J., Lowther, G. W. and Lee, W. R. (1994). Congenital ocular defects associated with an abnormality of the human chromosome 1: trisomy 1q32-qter. *J Pediatr Ophthalmol Strabismus* **31**, 41-5.

Costell, M., Carmona, R., Gustafsson, E., Gonzalez-Iriarte, M., Fassler, R. and Munoz-Chapuli, R. (2002). Hyperplastic conotruncal endocardial cushions and transposition of great arteries in perlecan-null mice. *Circ Res* **91**, 158-64.

Cripe, L., Andelfinger, G., Martin, L. J., Shooner, K. and Benson, D. W. (2004). Bicuspid aortic valve is heritable. *J Am Coll Cardiol* **44**, 138-43.

Crossin, K. L. and Hoffman, S. (1991). Expression of adhesion molecules during the formation and differentiation of the avian endocardial cushion tissue. *Dev Biol* **145**, 277-86.

Curran, M. E., Atkinson, D. L., Ewart, A. K., Morris, C. A., Leppert, M. F. and Keating, M. T. (1993). The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis. *Cell* **73**, 159-68.

de Jong, F., Opthof, T., Wilde, A. A., Janse, M. J., Charles, R., Lamers, W. H. and Moorman, A. F. (1992). Persisting zones of slow impulse conduction in developing chicken hearts. *Circ Res* **71**, 240-50.

De la Cruz, M. V., Gimenez-Ribotta, M., Saravalli, O. and Cayre, R. (1983). The contribution of the inferior endocardial cushion of the atrioventricular canal to cardiac septation and to the development of the atrioventricular valves: study in the chick embryo. *Am J Anat* **166**, 63-72.

de la Cruz, M. V., Sanchez Gomez, C., Arteaga, M. M. and Arguello, C. (1977). Experimental study of the development of the truncus and the conus in the chick embryo. *J Anat* **123**, 661-86.

De Vries, B. B., Knight, S. J., Homfray, T., Smithson, S. F., Flint, J. and Winter, R. M. (2001). Submicroscopic subtelomeric 1qter deletions: a recognisable phenotype? *J Med Genet* **38**, 175-8.

Delorme, B., Dahl, E., Jarry-Guichard, T., Briand, J. P., Willecke, K., Gros, D. and Theveniau-Ruissy, M. (1997). Expression pattern of connexin gene products at

the early developmental stages of the mouse cardiovascular system. *Circ Res* **81**, 423-37.

Diaz, G. A., Khan, K. T. and Gelb, B. D. (1998). The autosomal recessive Kenny-Caffey syndrome locus maps to chromosome 1q42-q43. *Genomics* **54**, 13-8.

Dickson, M. C., Slager, H. G., Duffie, E., Mummery, C. L. and Akhurst, R. J. (1993). RNA and protein localisations of TGF beta 2 in the early mouse embryo suggest an involvement in cardiac development. *Development* **117**, 625-39.

Dietz, H. C., Cutting, G. R., Pyeritz, R. E., Maslen, C. L., Sakai, L. Y., Corson, G. M., Puffenberger, E. G., Hamosh, A., Nanthakumar, E. J., Curristin, S. M. et al. (1991a). Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature* **352**, 337-9.

Dietz, H. C., Pyeritz, R. E., Hall, B. D., Cadle, R. G., Hamosh, A., Schwartz, J., Meyers, D. A. and Francomano, C. A. (1991b). The Marfan syndrome locus: confirmation of assignment to chromosome 15 and identification of tightly linked markers at 15q15-q21.3. *Genomics* **9**, 355-61.

Doliana, R., Bot, S., Mungiguerra, G., Canton, A., Cilli, S. P. and Colombatti, A. (2001). Isolation and characterization of EMILIN-2, a new component of the growing EMILINs family and a member of the EMI domain-containing superfamily. *J Biol Chem* **276**, 12003-11.

Duba, H. C., Erdel, M., Loffler, J., Bereuther, L., Fischer, H., Utermann, B. and Utermann, G. (1997). Detection of a de novo duplication of 1q32-qter by fluorescence in situ hybridisation in a boy with multiple malformations: further delineation of the trisomy 1q syndrome. *J Med Genet* **34**, 309-13.

Durocher, D., Charron, F., Warren, R., Schwartz, R. J. and Nemer, M. (1997). The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *Embo J* **16**, 5687-96.

Ehrman, L. A. and Yutzey, K. E. (1999). Lack of regulation in the heart forming region of avian embryos. *Dev Biol* **207**, 163-75.

Eisenberg, L. M. and Markwald, R. R. (1995). Molecular regulation of atrioventricular valvuloseptal morphogenesis. *Circ Res* **77**, 1-6.

Eisenberg, R., Young, D., Jacobson, B. and Boito, A. (1964). Familial Supravalvular Aortic Stenosis. *Am J Dis Child* **108**, 341-7.

Ekblom, P., Ekblom, M., Fecker, L., Klein, G., Zhang, H. Y., Kadoya, Y., Chu, M. L., Mayer, U. and Timpl, R. (1994). Role of mesenchymal nidogen for epithelial morphogenesis in vitro. *Development* **120**, 2003-14.

Elliott, D. A., Kirk, E. P., Yeoh, T., Chandar, S., McKenzie, F., Taylor, P., Grossfeld, P., Fatkin, D., Jones, O., Hayes, P. et al. (2003). Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol* **41**, 2072-6.

Ewart, A. K., Morris, C. A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., Stock, A. D., Leppert, M. and Keating, M. T. (1993a). Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. *Nat Genet* **5**, 11-6.

Ewart, A. K., Morris, C. A., Ensing, G. J., Loker, J., Moore, C., Leppert, M. and Keating, M. (1993b). A human vascular disorder, supravalvular aortic stenosis, maps to chromosome 7. *Proc Natl Acad Sci U S A* **90**, 3226-30.

Faustino, N. A. and Cooper, T. A. (2003). Pre-mRNA splicing and human disease. *Genes Dev* **17**, 419-37.

Ferencz, C., Neill, C. A., Boughman, J. A., Rubin, J. D., Brenner, J. I. and Perry, L. W. (1989). Congenital cardiovascular malformations associated with chromosome abnormalities: an epidemiologic study. *J Pediatr* **114**, 79-86.

Ferencz, C., Rubin, J. D., McCarter, R. J. and Clark, E. B. (1990). Maternal diabetes and cardiovascular malformations: predominance of double outlet right ventricle and truncus arteriosus. *Teratology* **41**, 319-26.

Fiegler, H., Carr, P., Douglas, E. J., Burford, D. C., Hunt, S., Smith, J., Vetrie, D., Gorman, P., Tomlinson, I. P. and Carter, N. P. (2003a). DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* **36**, 361-74.

Fiegler, H., Gribble, S. M., Burford, D. C., Carr, P., Prigmore, E., Porter, K. M., Clegg, S., Crolla, J. A., Dennis, N. R., Jacobs, P. et al. (2003b). Array painting: a method for the rapid analysis of aberrant chromosomes using DNA microarrays. *J Med Genet* **40**, 664-70.

Fox, J. W., Mayer, U., Nischt, R., Aumailley, M., Reinhardt, D., Wiedemann, H., Mann, K., Timpl, R., Krieg, T., Engel, J. et al. (1991). Recombinant nidogen consists of three globular domains and mediates binding of laminin to collagen type IV. *Embo J* **10**, 3137-46.

Galvin, K. M., Donovan, M. J., Lynch, C. A., Meyer, R. I., Paul, R. J., Lorenz, J. N., Fairchild-Huntress, V., Dixon, K. L., Dunmore, J. H., Gimbrone, M. A., Jr. et al. (2000). A role for smad6 in development and homeostasis of the cardiovascular system. *Nat Genet* **24**, 171-4.

Garani, G. P., Tamisari, L., Volpato, S. and Vigi, V. (1988). Terminal deletion of chromosome 1(q43) in a female infant. *J Med Genet* **25**, 211-2.

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

Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., Rothrock, C. R., Eapen, R. S., Hirayama-Yamada, K., Joo, K. et al. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* **424**, 443-7.

Gentile, M., Di Carlo, A., Volpe, P., Pansini, A., Nanna, P., Valenzano, M. C. and Buonadonna, A. L. (2003). FISH and cytogenetic characterization of a terminal chromosome 1q deletion: clinical case report and phenotypic implications. *Am J Med Genet* **117A**, 251-4.

George, E. L., Baldwin, H. S. and Hynes, R. O. (1997). Fibronectins are essential for heart and blood vessel morphogenesis but are dispensable for initial specification of precursor cells. *Blood* **90**, 3073-81.

Gill, H. K., Splitt, M., Sharland, G. K. and Simpson, J. M. (2003). Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J Am Coll Cardiol* **42**, 923-9.

Goldmuntz, E., Geiger, E. and Benson, D. W. (2001). NKX2.5 mutations in patients with tetralogy of fallot. *Circulation* **104**, 2565-8.

Gong, W., Gottlieb, S., Collins, J., Blescia, A., Dietz, H., Goldmuntz, E., McDonald-McGinn, D. M., Zackai, E. H., Emanuel, B. S., Driscoll, D. A. et al. (2001). Mutation analysis of TBX1 in non-deleted patients with features of DGS/VCFS or isolated cardiovascular defects. *J Med Genet* **38**, E45.

Green, E. K., Priestley, M. D., Waters, J., Maliszewska, C., Latif, F. and Maher, E. R. (2000). Detailed mapping of a congenital heart disease gene in chromosome 3p25. *J Med Genet* **37**, 581-7.

Gribble, S. M., Fiegler, H., Burford, D. C., Prigmore, E., Yang, F., Carr, P., Ng, B. L., Sun, T., Kamberov, E. S., Makarov, V. L. et al. (2004). Applications of combined DNA microarray and chromosome sorting technologies. *Chromosome Res* **12**, 35-43.

Hacia, J. G. (1999). Resequencing and mutational analysis using oligonucleotide microarrays. *Nat Genet* **21**, 42-7.

Hacia, J. G., Brody, L. C., Chee, M. S., Fodor, S. P. and Collins, F. S. (1996). Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nat Genet* **14**, 441-7.

Heilstedt, H. A., Ballif, B. C., Howard, L. A., Kashork, C. D. and Shaffer, L. G. (2003a). Population data suggest that deletions of 1p36 are a relatively common chromosome abnormality. *Clin Genet* **64**, 310-6.

Heilstedt, H. A., Ballif, B. C., Howard, L. A., Lewis, R. A., Stal, S., Kashork, C. D., Bacino, C. A., Shapira, S. K. and Shaffer, L. G. (2003b). Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. *Am J Hum Genet* **72**, 1200-12.

Henderson, D. J. and Copp, A. J. (1998). Versican expression is associated with chamber specification, septation, and valvulogenesis in the developing mouse heart. *Circ Res* **83**, 523-32.

Hoffman, J. I. and Kaplan, S. (2002). The incidence of congenital heart disease. *J Am Coll Cardiol* **39**, 1890-900.

Hopf, M., Gohring, W., Ries, A., Timpl, R. and Hohenester, E. (2001). Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1. *Nat Struct Biol* **8**, 634-40.

Ishkanian, A. S., Malloff, C. A., Watson, S. K., DeLeeuw, R. J., Chi, B., Coe, B. P., Snijders, A., Albertson, D. G., Pinkel, D., Marra, M. A. et al. (2004). A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* **36**, 299-303.

Jiang, X., Rowitch, D. H., Soriano, P., McMahon, A. P. and Sucov, H. M. (2000). Fate of the mammalian cardiac neural crest. *Development* **127**, 1607-16.

Johnson, V. P., Heck, L. J., Carter, G. A. and Flom, J. O. (1985). Deletion of the distal long arm of chromosome 1: a definable syndrome. *Am J Med Genet* **22**, 685-94.

Jones, C., Mullenbach, R., Grossfeld, P., Auer, R., Favier, R., Chien, K., James, M., Tunnacliffe, A. and Cotter, F. (2000). Co-localisation of CCG repeats and chromosome deletion breakpoints in Jacobsen syndrome: evidence for a common mechanism of chromosome breakage. *Hum Mol Genet* **9**, 1201-8.

Juberg, R. C., Haney, N. R. and Stallard, R. (1981). New deletion syndrome: 1q43. *Am J Hum Genet* **33**, 455-63.

Kelly, R. G., Brown, N. A. and Buckingham, M. E. (2001). The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm. *Dev Cell* **1**, 435-40.

Kelly, R. G. and Buckingham, M. E. (2002). The anterior heart-forming field: voyage to the arterial pole of the heart. *Trends Genet* **18**, 210-6.

Kessel, E., Pfeiffer, R. A., Blanke, W. and Schwarz, J. (1978). Terminal deletion of the long arm of chromosome 1 in a malformed newborn. *Hum Genet* **42**, 333-7.

Kim, J. S., Viragh, S., Moorman, A. F., Anderson, R. H. and Lamers, W. H. (2001). Development of the myocardium of the atrioventricular canal and the vestibular spine in the human heart. *Circ Res* **88**, 395-402.

Kirby, M. L., Gale, T. F. and Stewart, D. E. (1983). Neural crest cells contribute to normal aorticopulmonary septation. *Science* **220**, 1059-61.

Kitten, G. T., Markwald, R. R. and Bolender, D. L. (1987). Distribution of basement membrane antigens in cryopreserved early embryonic hearts. *Anat Rec* **217**, 379-90.

Kohfeldt, E., Sasaki, T., Gohring, W. and Timpl, R. (1998). Nidogen-2: a new basement membrane protein with diverse binding properties. *J Mol Biol* **282**, 99-109.

Krawczak, M., Reiss, J. and Cooper, D. N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet* **90**, 41-54.

Lamers, W. H., Viragh, S., Wessels, A., Moorman, A. F. and Anderson, R. H. (1995). Formation of the tricuspid valve in the human heart. *Circulation* **91**, 111-21.

Le Lievre, C. S. and Le Douarin, N. M. (1975). Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos. *J Embryol Exp Morphol* **34**, 125-54.

Levin, M., Johnson, R. L., Stern, C. D., Kuehn, M. and Tabin, C. (1995). A molecular pathway determining left-right asymmetry in chick embryogenesis. *Cell* **82**, 803-14.

Levy, H. L., Guldberg, P., Guttler, F., Hanley, W. B., Matalon, R., Rouse, B. M., Trefz, F., Azen, C., Allred, E. N., de la Cruz, F. et al. (2001). Congenital heart disease in maternal phenylketonuria: report from the Maternal PKU Collaborative Study. *Pediatr Res* **49**, 636-42.

Li, D. Y., Toland, A. E., Boak, B. B., Atkinson, D. L., Ensing, G. J., Morris, C. A. and Keating, M. T. (1997a). Elastin point mutations cause an obstructive vascular disease, supravalvular aortic stenosis. *Hum Mol Genet* **6**, 1021-8.

Li, Q. Y., Newbury-Ecob, R. A., Terrett, J. A., Wilson, D. I., Curtis, A. R., Yi, C. H., Gebuhr, T., Bullen, P. J., Robson, S. C., Strachan, T. et al. (1997b). Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat Genet* **15**, 21-9.

Lichtner, P., Attie-Bitach, T., Schuffenhauer, S., Henwood, J., Bouvagnet, P., Scambler, P. J., Meitinger, T. and Vekemans, M. (2002). Expression and mutation analysis of BRUNOL3, a candidate gene for heart and thymus developmental defects associated with partial monosomy 10p. *J Mol Med* **80**, 431-42.

Lin, Q., Srivastava, D. and Olson, E. N. (1997). A transcriptional pathway for cardiac development. *Cold Spring Harb Symp Quant Biol* **62**, 405-11.

Lindsay, E. A., Vitelli, F., Su, H., Morishima, M., Huynh, T., Pramparo, T., Jurecic, V., Ogunrinu, G., Sutherland, H. F., Scambler, P. J. et al. (2001). Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* **410**, 97-101.

Logan, M., Pagan-Westphal, S. M., Smith, D. M., Paganessi, L. and Tabin, C. J. (1998). The transcription factor Pitx2 mediates situs-specific morphogenesis in response to left-right asymmetric signals. *Cell* **94**, 307-17.

Lough, J. and Sugi, Y. (2000). Endoderm and heart development. *Dev Dyn* **217**, 327-42.

Lyons, I., Parsons, L. M., Hartley, L., Li, R., Andrews, J. E., Robb, L. and Harvey, R. P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev* **9**, 1654-66.

Magenis, R. E., Maslen, C. L., Smith, L., Allen, L. and Sakai, L. Y. (1991). Localization of the fibrillin (FBN) gene to chromosome 15, band q21.1. *Genomics* **11**, 346-51.

Maitra, A., Cohen, Y., Gillespie, S. E., Mambo, E., Fukushima, N., Hoque, M. O., Shah, N., Goggins, M., Califano, J., Sidransky, D. et al. (2004). The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. *Genome Res* **14**, 812-9.

Mankinen, C. B., Sears, J. W. and Alvarez, V. R. (1976). Terminal (1)(q43) long-arm deletion of chromosome no. 1 in a three-year-old female. *Birth Defects Orig Artic Ser* **12**, 131-6.

Manouvrier-Hanu, S., Walbaum, R. and Gayot, C. (1986). A new case of distal deletion of the long arm of chromosome 1. *Am J Med Genet* **25**, 599-600.

Mantripragada, K. K., Buckley, P. G., Jarbo, C., Menzel, U. and Dumanski, J. P. (2003). Development of NF2 gene specific, strictly sequence defined diagnostic microarray for deletion detection. *J Mol Med* **81**, 443-51.

Mantripragada, K. K., Tapia-Paez, I., Blennow, E., Nilsson, P., Wedell, A. and Dumanski, J. P. (2004). DNA copy-number analysis of the 22q11 deletion-syndrome region using array-CGH with genomic and PCR-based targets. *Int J Mol Med* **13**, 273-9.

Marino, B. and Digilio, M. C. (2000). Congenital heart disease and genetic syndromes: specific correlation between cardiac phenotype and genotype. *Cardiovasc Pathol* **9**, 303-15.

Markwald, R. R., Krook, J. M., Kitten, G. T. and Runyan, R. B. (1981). Endocardial cushion tissue development: structural analyses on the attachment of extracellular matrix to migrating mesenchymal cell surfaces. *Scan Electron Microsc*, 261-74.

Markwald, R. R., Mjaatvedt, C. H., Krug, E. L. and Sinning, A. R. (1990). Inductive interactions in heart development. Role of cardiac adherons in cushion tissue formation. *Ann N Y Acad Sci* **588**, 13-25.

Marvin, M. J., Di Rocco, G., Gardiner, A., Bush, S. M. and Lassar, A. B. (2001). Inhibition of Wnt activity induces heart formation from posterior mesoderm. *Genes Dev* **15**, 316-27.

Mayer, U., Nischt, R., Poschl, E., Mann, K., Fukuda, K., Gerl, M., Yamada, Y. and Timpl, R. (1993). A single EGF-like motif of laminin is responsible for high affinity nidogen binding. *Embo J* **12**, 1879-85.

McDermid, H. E. and Morrow, B. E. (2002). Genomic disorders on 22q11. *Am J Hum Genet* **70**, 1077-88.

Meinecke, P. and Vogtel, D. (1987). A specific syndrome due to deletion of the distal long arm of chromosome 1. *Am J Med Genet* **28**, 371-6.

Merlob, P., Kohn, G., Litwin, A., Nissenkorn, I., Katznelson, M. B. and Reisner, S. H. (1989). New chromosome aberration: duplication of a large part of chromosome 4q and partial deletion of chromosome 1q. *Am J Med Genet* **32**, 22-6.

Merscher, S., Funke, B., Epstein, J. A., Heyer, J., Puech, A., Lu, M. M., Xavier, R. J., Demay, M. B., Russell, R. G., Factor, S. et al. (2001). TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* **104**, 619-29.

Millar, J. K., Wilson-Annan, J. C., Anderson, S., Christie, S., Taylor, M. S., Semple, C. A., Devon, R. S., Clair, D. M., Muir, W. J., Blackwood, D. H. et al. (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum Mol Genet* **9**, 1415-23.

Mitchell, S. C., Sellmann, A. H., Westphal, M. C. and Park, J. (1971). Etiologic correlates in a study of congenital heart disease in 56,109 births. *Am J Cardiol* **28**, 653-7.

Mjaatvedt, C. H., Krug, E. L. and Markwald, R. R. (1991). An antiserum (ES1) against a particulate form of extracellular matrix blocks the transition of cardiac endothelium into mesenchyme in culture. *Dev Biol* **145**, 219-30.

Mjaatvedt, C. H., Nakaoka, T., Moreno-Rodriguez, R., Norris, R. A., Kern, M. J., Eisenberg, C. A., Turner, D. and Markwald, R. R. (2001). The outflow tract of the heart is recruited from a novel heart-forming field. *Dev Biol* **238**, 97-109.

Mjaatvedt, C. H., Yamamura, H., Capehart, A. A., Turner, D. and Markwald, R. R. (1998). The Cspg2 gene, disrupted in the hdf mutant, is required for right cardiac chamber and endocardial cushion formation. *Dev Biol* **202**, 56-66.

Molkentin, J. D., Lin, Q., Duncan, S. A. and Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev* **11**, 1061-72.

Moorman, A. F., de Jong, F., Denyn, M. M. and Lamers, W. H. (1998). Development of the cardiac conduction system. *Circ Res* **82**, 629-44.

Murayama, K., Greenwood, R. S., Rao, K. W. and Aylsworth, A. S. (1991). Neurological aspects of del(1q) syndrome. *Am J Med Genet* **40**, 488-92.

Murshed, M., Smyth, N., Miosge, N., Karolat, J., Krieg, T., Paulsson, M. and Nischt, R. (2000). The absence of nidogen 1 does not affect murine basement membrane formation. *Mol Cell Biol* **20**, 7007-12.

Nakajima, Y., Krug, E. L. and Markwald, R. R. (1994). Myocardial regulation of transforming growth factor-beta expression by outflow tract endothelium in the early embryonic chick heart. *Dev Biol* **165**, 615-26.

Nakajima, Y., Morishima, M., Nakazawa, M., Momma, K. and Nakamura, H. (1997). Distribution of fibronectin, type I collagen, type IV collagen, and laminin in the cardiac jelly of the mouse embryonic heart with retinoic acid-induced complete transposition of the great arteries. *Anat Rec* **249**, 478-85.

Nakamura, T., Lozano, P. R., Ikeda, Y., Iwanaga, Y., Hinek, A., Minamisawa, S., Cheng, C. F., Kobuke, K., Dalton, N., Takada, Y. et al. (2002). Fibulin-5/DANCE is essential for elastogenesis in vivo. *Nature* **415**, 171-5.

Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., Kido, M. and 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**, 829-37.

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

Nowaczyk, M. J., Bayani, J., Freeman, V., Watts, J., Squire, J. and Xu, J. (2003). De novo 1q32q44 duplication and distal 1q trisomy syndrome. *Am J Med Genet* **120A**, 229-33.

Olin, A. I., Morgelin, M., Sasaki, T., Timpl, R., Heinegard, D. and Aspberg, A. (2001). The proteoglycans aggrecan and Versican form networks with fibulin-2 through their lectin domain binding. *J Biol Chem* **276**, 1253-61.

Packham, E. A. and Brook, J. D. (2003). T-box genes in human disorders. *Hum Mol Genet* **12 Spec No 1**, R37-44.

Parvari, R., Hershkovitz, E., Grossman, N., Gorodischer, R., Loeys, B., Zecic, A., Mortier, G., Gregory, S., Sharony, R., Kambouris, M. et al. (2002). Mutation of TBCE causes hypoparathyroidism-retardation-dysmorphism and autosomal recessive Kenny-Caffey syndrome. *Nat Genet* **32**, 448-52.

Parvari, R., Hershkovitz, E., Kanis, A., Gorodischer, R., Shalitin, S., Sheffield, V. C. and Carmi, R. (1998). Homozygosity and linkage-disequilibrium mapping of the syndrome of congenital hypoparathyroidism, growth and mental retardation, and dysmorphism to a 1-cM interval on chromosome 1q42-43. *Am J Hum Genet* **63**, 163-9.

Pelliccia, F., Limongi, M. Z., Gaddini, L. and Rocchi, A. (1998). Assignment of FRA1H common fragile site to human chromosome band 1q42.1 proximal to the nuclear NAD+ ADP-ribosyltransferase gene (ADPRT) and to the main 5S rRNA gene locus. *Cytogenet Cell Genet* **82**, 121-2.

Penny, L. A., Dell'Aquila, M., Jones, M. C., Bergoffen, J., Cunniff, C., Fryns, J. P., Grace, E., Graham, J. M., Jr., Kousseff, B., Mattina, T. et al. (1995). Clinical and molecular characterization of patients with distal 11q deletions. *Am J Hum Genet* **56**, 676-83.

Pfeifer, D., Kist, R., Dewar, K., Devon, K., Lander, E. S., Birren, B., Korniszewski, L., Back, E. and Scherer, G. (1999). Campomelic dysplasia

translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. *Am J Hum Genet* **65**, 111-24.

Phillips, H. M., Renforth, G. L., Spalluto, C., Hearn, T., Curtis, A. R., Craven, L., Havarani, B., Clement-Jones, M., English, C., Stumper, O. et al. (2002). Narrowing the critical region within 11q24-qter for hypoplastic left heart and identification of a candidate gene, JAM3, expressed during cardiogenesis. *Genomics* **79**, 475-8.

Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y. et al. (1998). High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* **20**, 207-11.

Poelmann, R. E., Mikawa, T. and Gittenberger-de Groot, A. C. (1998). Neural crest cells in outflow tract septation of the embryonic chicken heart: differentiation and apoptosis. *Dev Dyn* **212**, 373-84.

Poschl, E., Schlotzer-Schrehardt, U., Brachvogel, B., Saito, K., Ninomiya, Y. and Mayer, U. (2004). Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development* **131**, 1619-28.

Putnam, E. A., Zhang, H., Ramirez, F. and Milewicz, D. M. (1995). Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder, congenital contractural arachnodactyly. *Nat Genet* **11**, 456-8.

Ranger, A. M., Grusby, M. J., Hodge, M. R., Gravallese, E. M., de la Brousse, F. C., Hoey, T., Mickanin, C., Baldwin, H. S. and Glimcher, L. H. (1998). The transcription factor NF-ATc is essential for cardiac valve formation. *Nature* **392**, 186-90.

Rauen, K. A., Albertson, D. G., Pinkel, D. and Cotter, P. D. (2002). Additional patient with del(12)(q21.2q22): further evidence for a candidate region for cardio-facio-cutaneous syndrome? *Am J Med Genet* **110**, 51-6.

Reinhardt, D. P., Sasaki, T., Dzamba, B. J., Keene, D. R., Chu, M. L., Gohring, W., Timpl, R. and Sakai, L. Y. (1996). Fibrillin-1 and fibulin-2 interact and are colocalized in some tissues. *J Biol Chem* **271**, 19489-96.

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

Richard, I. and Beckmann, J. S. (1995). How neutral are synonymous codon mutations? *Nat Genet* **10**, 259.

Runyan, R. B., Potts, J. D. and Weeks, D. L. (1992). TGF-beta 3-mediated tissue interaction during embryonic heart development. *Mol Reprod Dev* **32**, 152-9.

Salmivirta, K., Talts, J. F., Olsson, M., Sasaki, T., Timpl, R. and Ekblom, P. (2002). Binding of mouse nidogen-2 to basement membrane components and cells and its expression in embryonic and adult tissues suggest complementary functions of the two nidogens. *Exp Cell Res* **279**, 188-201.

Sarda, P., Lefort, G., Taviaux, S., Humeau, C. and Rieu, D. (1992). Interstitial deletion of chromosome 1 del (1) (q32 q42): case report and review of the literature. *Clin Genet* **41**, 25-7.

Sasaki, T., Gohring, W., Miosge, N., Abrams, W. R., Rosenbloom, J. and Timpl, R. (1999). Tropoelastin binding to fibulins, nidogen-2 and other extracellular matrix proteins. *FEBS Lett* **460**, 280-4.

Sasaki, T., Gohring, W., Pan, T. C., Chu, M. L. and Timpl, R. (1995a). Binding of mouse and human fibulin-2 to extracellular matrix ligands. *J Mol Biol* **254**, 892-9.

Sasaki, T., Kostka, G., Gohring, W., Wiedemann, H., Mann, K., Chu, M. L. and Timpl, R. (1995b). Structural characterization of two variants of fibulin-1 that differ in nidogen affinity. *J Mol Biol* **245**, 241-50.

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

Schoenwolf, G. C. and Garcia-Martinez, V. (1995). Primitive-streak origin and state of commitment of cells of the cardiovascular system in avian and mammalian embryos. *Cell Mol Biol Res* **41**, 233-40.

Schott, J. J., Benson, D. W., Basson, C. T., Pease, W., Silberbach, G. M., Moak, J. P., Maron, B. J., Seidman, C. E. and Seidman, J. G. (1998). Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* **281**, 108-11.

Schroeder, J. A., Jackson, L. F., Lee, D. C. and Camenisch, T. D. (2003). Form and function of developing heart valves: coordination by extracellular matrix and growth factor signaling. *J Mol Med* **81**, 392-403.

Schuffenhauer, S., Lichtner, P., Peykar-Derakhshandeh, P., Murken, J., Haas, O. A., Back, E., Wolff, G., Zabel, B., Barisic, I., Rauch, A. et al. (1998). Deletion mapping on chromosome 10p and definition of a critical region for the second DiGeorge syndrome locus (DGS2). *Eur J Hum Genet* **6**, 213-25.

Schultheiss, T. M., Burch, J. B. and Lassar, A. B. (1997). A role for bone morphogenetic proteins in the induction of cardiac myogenesis. *Genes Dev* **11**, 451-62.

Schultheiss, T. M., Xydas, S. and Lassar, A. B. (1995). Induction of avian cardiac myogenesis by anterior endoderm. *Development* **121**, 4203-14.

Sedmera, D., Pexieder, T., Rychterova, V., Hu, N. and Clark, E. B. (1999). Remodeling of chick embryonic ventricular myoarchitecture under experimentally changed loading conditions. *Anat Rec* **254**, 238-52.

Snijders, A. M., Nowak, N., Segraves, R., Blackwood, S., Brown, N., Conroy, J., Hamilton, G., Hindle, A. K., Huey, B., Kimura, K. et al. (2001). Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* **29**, 263-4.

Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nickolenko, J., Benner, A., Dohner, H., Cremer, T. and Lichter, P. (1997). Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* **20**, 399-407.

Spence, S. G., Argraves, W. S., Walters, L., Hungerford, J. E. and Little, C. D. (1992). Fibulin is localized at sites of epithelial-mesenchymal transitions in the early avian embryo. *Dev Biol* **151**, 473-84.

Srivastava, D. and Olson, E. N. (2000). A genetic blueprint for cardiac development. *Nature* **407**, 221-6.

Srivastava, D., Thomas, T., Lin, Q., Kirby, M. L., Brown, D. and Olson, E. N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nat Genet* **16**, 154-60.

St Clair, D., Blackwood, D., Muir, W., Carothers, A., Walker, M., Spowart, G., Gosden, C. and Evans, H. J. (1990). Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* **336**, 13-6.

Steffensen, D. M., Chu, E. H., Speert, D. P., Wall, P. M., Meilinger, K. and Kelch, R. P. (1977). Partial trisomy of the long arm of human chromosome 1 as demostrated by in situ hybridization with 5S ribosomal RNA. *Hum Genet* **36**, 25-33.

Supp, D. M., Witte, D. P., Potter, S. S. and Brueckner, M. (1997). Mutation of an axonemal dynein affects left-right asymmetry in inversus viscerum mice. *Nature* **389**, 963-6.

Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N. and Izumo, S. (1999). The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development* **126**, 1269-80.

Telenius, H., Carter, N. P., Bebb, C. E., Nordenskjold, M., Ponder, B. A. and Tunnacliffe, A. (1992). Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* **13**, 718-25.

Tennstedt, C., Chaoui, R., Korner, H. and Dietel, M. (1999). Spectrum of congenital heart defects and extracardiac malformations associated with chromosomal abnormalities: results of a seven year necropsy study. *Heart* **82**, 34-9.

Terrett, J. A., Newbury-Ecob, R., Cross, G. S., Fenton, I., Raeburn, J. A., Young, I. D. and Brook, J. D. (1994). Holt-Oram syndrome is a genetically heterogeneous disease with one locus mapping to human chromosome 12q. *Nat Genet* **6**, 401-4.

Timpl, R. and Brown, J. C. (1996). Supramolecular assembly of basement membranes. *Bioessays* **18**, 123-32.

Timpl, R., Sasaki, T., Kostka, G. and Chu, M. L. (2003). Fibulins: a versatile family of extracellular matrix proteins. *Nat Rev Mol Cell Biol* **4**, 479-89.

Tiso, N., Stephan, D. A., Nava, A., Bagattin, A., Devaney, J. M., Stanchi, F., Larderet, G., Brahmbhatt, B., Brown, K., Bauce, B. et al. (2001). Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythmogenic right ventricular cardiomyopathy type 2 (ARVD2). *Hum Mol Genet* **10**, 189-94.

Tsuda, T., Wang, H., Timpl, R. and Chu, M. L. (2001). Fibulin-2 expression marks transformed mesenchymal cells in developing cardiac valves, aortic arch vessels, and coronary vessels. *Dev Dyn* **222**, 89-100.

Van Esch, H., Groenen, P., Nesbit, M. A., Schuffenhauer, S., Lichtner, P., Vanderlinden, G., Harding, B., Beetz, R., Bilous, R. W., Holdaway, I. et al.

(2000). GATA3 haplo-insufficiency causes human HDR syndrome. *Nature* **406**, 419-22.

Viragh, S. and Challice, C. E. (1973). Origin and differentiation of cardiac muscle cells in the mouse. *J Ultrastruct Res* **42**, 1-24.

Vissers, L. E., van Ravenswaaij, C. M., Admiraal, R., Hurst, J. A., de Vries, B. B., Janssen, I. M., van der Vliet, W. A., Huys, E. H., de Jong, P. J., Hamel, B. C. et al. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* **36**, 955-7.

Waldo, K. L., Kumiski, D. H., Wallis, K. T., Stadt, H. A., Hutson, M. R., Platt, D. H. and Kirby, M. L. (2001). Conotruncal myocardium arises from a secondary heart field. *Development* **128**, 3179-88.

Watanabe, M., Choudhry, A., Berlan, M., Singal, A., Siwik, E., Mohr, S. and Fisher, S. A. (1998). Developmental remodeling and shortening of the cardiac outflow tract involves myocyte programmed cell death. *Development* **125**, 3809-20.

Watkins, M. L., Rasmussen, S. A., Honein, M. A., Botto, L. D. and Moore, C. A. (2003). Maternal obesity and risk for birth defects. *Pediatrics* **111**, 1152-8.

Watson, M. S., Gargus, J. J., Blakemore, K. J., Katz, S. N. and Breg, W. R. (1986). Chromosome deletion 1q42-43. *Am J Med Genet* **24**, 1-6.

Webb, S., Kanani, M., Anderson, R. H., Richardson, M. K. and Brown, N. A. (2001). Development of the human pulmonary vein and its incorporation in the morphologically left atrium. *Cardiol Young* **11**, 632-42.

Wright, L. L., Schwartz, M. F., Schwartz, S. and Karesh, J. (1986). An unusual ocular finding associated with chromosome 1q deletion syndrome. *Pediatrics* **77**, 786.

Ya, J., van den Hoff, M. J., de Boer, P. A., Tesink-Taekema, S., Franco, D., Moorman, A. F. and Lamers, W. H. (1998). Normal development of the outflow tract in the rat. *Circ Res* **82**, 464-72.

Yarnitzky, T. and Volk, T. (1995). Laminin is required for heart, somatic muscles, and gut development in the Drosophila embryo. *Dev Biol* **169**, 609-18.

Yasui, H., Nakazawa, M., Morishima, M., Miyagawa-Tomita, S. and Momma, K. (1995). Morphological observations on the pathogenetic process of transposition of the great arteries induced by retinoic acid in mice. *Circulation* **91**, 2478-86.

Youssoufian, H., Chance, P., Tuck-Muller, C. M. and Jabs, E. W. (1988). Association of a new chromosomal deletion [del(1)(q32q42)] with diaphragmatic hernia: assignment of a human ferritin gene. *Hum Genet* **78**, 267-70.

Yu, W., Ballif, B. C., Kashork, C. D., Heilstedt, H. A., Howard, L. A., Cai, W. W., White, L. D., Liu, W., Beaudet, A. L., Bejjani, B. A. et al. (2003). Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions. *Hum Mol Genet* **12**, 2145-52.

Zanetti, M., Braghetta, P., Sabatelli, P., Mura, I., Doliana, R., Colombatti, A., Volpin, D., Bonaldo, P. and Bressan, G. M. (2004). EMILIN-1 deficiency induces elastogenesis and vascular cell defects. *Mol Cell Biol* **24**, 638-50.

Zhang, H. Y., Chu, M. L., Pan, T. C., Sasaki, T., Timpl, R. and Ekblom, P. (1995). Extracellular matrix protein fibulin-2 is expressed in the embryonic endocardial cushion tissue and is a prominent component of valves in adult heart. *Dev Biol* **167**, 18-26.

Zhu, X., Sasse, J., McAllister, D. and Lough, J. (1996). Evidence that fibroblast growth factors 1 and 4 participate in regulation of cardiogenesis. *Dev Dyn* **207**, 429-38.

Zimmermann, K., Hoischen, S., Hafner, M. and Nischt, R. (1995). Genomic sequences and structural organization of the human nidogen gene (NID). *Genomics* **27**, 245-50.