CHARACTERIZATION OF CIRRIN, A CANDIDATE FOR CONGENITAL HEART DEFECTS IN 3p- SYNDROME

by

Paul A. Rupp

A DISSERTATION

Presented to the Department of Molecular and Medical

Genetics and the Oregon Health Sciences

University School of Medicine

in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

December 1999

School of Medicine
Oregon Health Sciences University

CERTIFICATE OF APPROVAL

This is to certify that the Ph.D. thesis of

Paul A. Rupp

has been approved

[redacted]
Professor in charge of thesis

[redacted]
Member

[redacted]
Member

[redacted]
Member

[redacted]

Dean for Graduate Studies

Table of Contents

Figures and Tables

Abbreviations

Abbreviation	
AV	Atrioventricular
AVCD	Atrioventricular Canal Defect
BAC	Bacterial Artificial Chromosome
CAM	Cell Adhesion Molecule
CHD	Congenital Heart Defects
cM	Centimorgan
DEPC	Diethylpyrocarbonate
DS	Down Syndrome
ECD	Endocardial Cushion Defects
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
cb-EGF	Calcium Binding Epidermal Growth Factor
EST	Expressed Sequence Tag
FISH	Fluorescent *in situ* Hybridization
FN	Fibronectin
ES	EDTA Soluble
HH	Hamburger and Hamilton
hLAMP	Heart Lectin Associated Myocardial Protein
HOS	Holt-Oram Syndrome
kb	Kilobase
LR	Left-Right

MCCM	Mesenchyme Cushion Conditioned Medium
MCM	Myocardial Conditioned Medium
OT	Outflow Tract
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline + Tween-20
PMCA2	Plasma Membrane Calcium Transporting ATPase isoform 2
SAVC	Superior Atrioventricular Cushion
SVCC	Sinistro-ventral Conal Cushion
TGF-β	Transforming Growth Factor-β
TBS	Tris-buffered Saline
TBST	Tris-buffered Saline + Tween-20
UTR	Untranslated Region
VHL	Von Hippel-Lindau
WE	Tryptophan/Glutamic Acid rich

Acknowledgements

During my time at OHSU, a number of people have helped and supported me in getting to this point. I would like to take this time to acknowledge and thank them for all they have done throughout the years. I would like to begin by thanking my family (Mom, Pierre, Chris, Dad, Cathy, Jenny, Rebecca, "Gramps" and Sue) for all of their encouragement, love and support. I would not have accomplished this if it were not for you. Thanks.

Of course, I would not be here in the first place if it were not for my advisor, Dr. Cheryl Maslen. I know through the years I have not exactly been a stellar graduate student, but you have helped and supported me through many trying times. I want to thank you for sticking by me and not giving up on me. You have been a wonderful person to work for and I will miss you. You have taught me much that I may bring with me on my next exciting journey.

Darcie and Gameil, what can I say. I have spent more time with the two of you than any others. I have learned much from both of you and have truly appreciated our time together, both in lab and without. I will miss "escaping" from the hill, our little lab happy hours and just hanging out with the two of you.

I would like to give a special thanks to Dr. Robert Glanville, both a member of my committee and a collaborator of ours. Thank you for taking the time to show me the ropes with regards to protein work. I learned very much from you and have appreciated all of your help.

At this time I would like to thank others for work they have done that has contributed to this dissertation: Gamiel, for your work with cirrin in baculovirus; Huey-Ju Kuo and Noi Tran, for all your help with antibody work; Dr. Susan Olson and Carol Reifsteck for their hard work on 3p- FISH analysis; Drs. Cynthia Timmers and Markus Grompe for bringing the cirrin EST to our attention; Jim Hejna for all his help with the BAC; and Dr. Kent Thornburg for helping to interpret the developmental data.

I would also like to thank my committee members, current and former, for their time and guidance. Thank you Dr. Robert Glanville, Dr. Matt Thayer, Dr. Jon Zonana, Dr. Susan Olson, and Dr. Mike Forte. I appreciate all you have done for me.

Abstract

Partial deletion of the short arm of chromosome 3 (3p25-pter) gives rise to the human cytogenetic disorder 3p- syndrome. The phenotype includes dysmorphic facial features, microcephaly, and growth and mental retardation. In addition, approximately 1/3 of patients have cardiac septal defects. Molecular genetic analysis of the chromosome breakpoints have defined the critical region for cardiac malformations. Here is described the identification and characterization of a highly conserved gene encoding an extracellular protein named cirrin. High levels of cirrin mRNA expression are observed in the endocardial cushions and myocardium of the developing heart. The cirrin locus is at 3p25 and lies within the critical region for cardiac defects associated with 3p- syndrome, making it a compelling candidate gene for these heart malformations. Examination of 3p- cell lines that define the critical region shows that deletion of the cirrin gene correlates with the occurrence of congenital heart defects. The complete cDNA sequence of the cirrin gene, its genomic organization, protein domain structure, and pattern of expression, are presented here. Partial protein characterization demonstrates that cirrin is an extracellular protein with some features common to matrix proteins. However, lack of extensive similarity to other proteins indicates that it is not a member of any known protein family and its function is as yet unknown. In addition to its association with cardiac septal defects in 3p- syndrome, it is proposed that the cirrin gene is also a viable candidate for similar heart defects of unknown etiology.

Chapter 1

Introduction

1.) Cardiac Development:

In developing vertebrates, the heart is the first organ formed with the earliest cardiac structure evident, in humans, at 3 weeks of gestation (Srivastava 1999). Normal heart development progresses from bilateral heart forming fields into a primitive, tubular structure and ultimately into a functioning four-chambered heart. Alterations of any of the numerous steps involved in cardiogenesis are likely to result in cardiac defects. Appearing in nearly one percent of newborn infants (Eisenberg and Markwald 1995), congenital heart defects (CHD) are the most common form of birth defect and are the major cause of premature death associated with congenital abnormalities. The majority of CHD are caused by the improper formation of valves and the membranous septa in the developing heart (Potts, Dagle et al. 1991). Individuals born with a CHD represent a small fraction of the total number of patients with cardiac defects, with more severe cardiac abnormalities resulting in spontaneous abortions (Hoffman 1995). To understand normal heart development and the many routes to congenital heart defects, identification and functional characterization of genes expressed and proteins produced during cardiac development is key. What follows is a brief overview of cardiac development with specific emphasis placed on endocardial cushion formation and subsequent valvuloseptal morphogenesis.

In all vertebrates, the primary heart tube is created from the ventral midline fusion of two primordial heart-forming fields (Eisenberg and Markwald 1995) resulting in a hollow cylinder consisting of two concentric epithelial layers: the endocardium surrounded by the thicker myocardium. The two epithelial layers are separated by an acellular matrix traditionally referred to as cardiac jelly (Davis 1924). The fusion,

occurring anteriorly, forms a series of primitive tubular segments (Cruz, Sanchez-Gomez et al. 1989). Superficially, each of the five segments of the primary heart tube appears to be homogeneous throughout the anterior-posterior axis, however sub-populations of cells found in the atrioventricular (AV) canal and outflow tract (conotruncus) are functionally distinct in their ability to form endocardial cushion tissue (Figure 1) (Mjaatvedt, Yamamura et al. 1999).

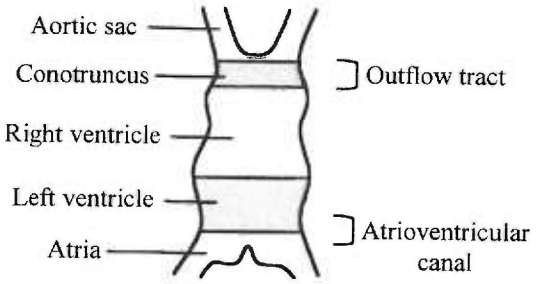


Figure 1. Diagram of five segments of linear heart with areas involved in endocardial cushion morphogenesis indicated on the right side.

As the heart forms, the cardiac jelly in the AV canal and proximal outflow tract (OT) expand predominantly through myocardial secretion of extracellular matrix (ECM) proteins such as laminin, proteoglycans, collagen, fibronectin, vitronectin and fibulin-1 (Kitten, Markwald et al. 1987; Bouchey, Argraves et al. 1996). The cardiac jelly exists as a fusion between a large myocardial derived basement membrane and a smaller endocardial basement membrane (Kitten, Markwald et al. 1987). The matrix formation is important for cell adhesion and migration of mesenchymal cells later in development. The mechanisms that regulate the secretion of matrix proteins into these regions are unknown.

Once the endocardial cushions have expanded, the myocardium secretes a particulate form of matrix referred to as adherons. The cardiac adherons, produced only in the AV canal and OT (Figure 1 and 2), induce a sub-population of endocardial cells to undergo a transformation to mesenchyme (Mjaatvedt, Yamamura et al. 1999). Adherons are aggregates of several ECM proteins, including fibronectin, transferrin, hLAMP-1, ES/130, as well as other ES (EDTA soluble) proteins (Kitten, Markwald et al. 1987; Rezaee, Isokawa et al. 1993; Isokawa, Rezaee et al. 1994; Sinning and Hewitt 1996). The particulate matrix can be removed by EDTA extraction from AV myocardium and is capable of inducing an epithelial-mesenchyme transition *in vitro* (Krug, Runyan et al. 1987). Antibodies raised against heart lectin-associated myocardial proteins (hLAMP) are capable of removing the inductive signals in culture assays (Sinning, Hewitt et al. 1995). Similarly, antibodies raised against ES aggregates, as a whole, are able to block the inductive activity in ES extracts as well as myocardial conditioned medium (Mjaatvedt, Krug et al. 1991). More specifically, antibodies and antisense oligonucleotides to the ECM molecule ES/130 have been shown to block the epithelial-mesenchymal transition (Mjaatvedt, Krug et al. 1991; Rezaee, Isokawa et al. 1993). ES/130 is expressed first in the myocardium and then in the endocardium-mesenchyme (Mjaatvedt, Yamamura et al. 1999). Endothelial-derived mesenchyme cells have been shown to migrate toward regions of increasing adheron density (near the myocardium) (Kitten, Markwald et al. 1987), thus populating the cardiac jelly.

The sub-population of endocardial cells capable of undergoing an epithelial to mesenchyme transition express the JB3 marker (fibrillin-2) (Rongish, Drake et al. 1998) and is found only in the AV/OT regions after Hamburger and Hamilton (HH) stages 11-

13 (staging reviewed in Table 1). Fibrillin-2 is an extracellular matrix protein found in both elastic and non-elastic connective tissues (Zhang, Apfelroth et al. 1994). In explant experiments on collagen gels, after mesenchyme transition, the JB3 marker is associated with the mesenchymal cells as well as being located within a network of fibrillar material (Wunsch, Little et al. 1994).

Table 1. Summary of stages of chick cardiac development.

	HH Stage	Hrs. Incubation	# Somites
Migration of precardiac cells	4	18	0
Assembly of myocardial plate	5	19-22	0
Generation of single heart tube	9	27-30	7
Tubular heart begins contractions	10	33-38	10
Looping begins	11	40-45	13
Endothelial-mesenchyme signal	14	44-52	20
Endothelial cells (EC) activated	16	50-56	27
EC separation and transformation	17	55-64	30
Mesenchyme cell (MC) invasion	18	72	36
Cushion tissue mesenchyme formed	17-20		
MC invasion continues	18-22		
Cushion remodeling	24-28		

Transforming growth factor-β (TGF-β) is another factor involved in the transformation of epithelial cells to mesenchyme. Addition of antibodies raised against TGF-β (no distinction as to β1-3) will inhibit both endothelial cell activation as well as mesenchyme cell invasion in chick embryo explants (Potts and Runyan 1989). Specifically inhibiting TGF-β3 with antibodies or antisense oligonucleotides blocks mesenchyme formation in chick hearts (Potts, Dagle et al. 1991; Nakajima, Krug et al. 1994). Ramsdell and coworkers demonstrated that TGF-β3 expression occurs in

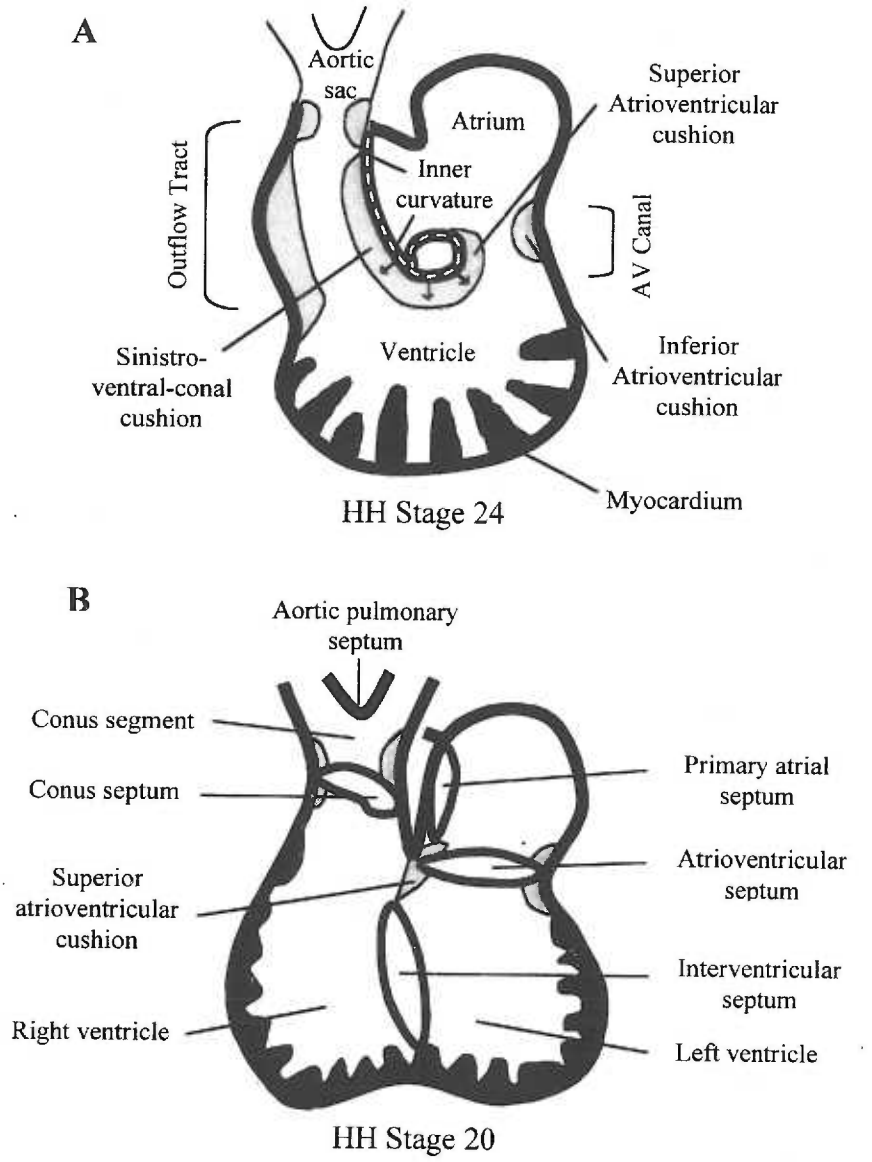


Figure 2. **A**, Segmental interaction during remodeling of the inner curvature. The inner curvature of the heart is completely lined by cushion formed by the fusion of the sinistro-ventral-conal cushion and the superior atrioventricular cushion. During myocardialization, myocardial cells from the inner curvature (dashed line) invade the cellularized cushions (arrows). **B**, Diagram depicting alignment of the septal ridges. Rings indicate future septum (Adapted from Mjaatvedt et. al, 1999). Light gray areas are cushion tissues. Heavy dark lines are myocardium.

transforming cells as a response to a myocardial induction signal (Ramsdell and Markwald 1997). When myocardial conditioned medium (MCM) and mesenchyme cushion conditioned medium (MCCM) are immunoadsorbed with TGF-β3 antibodies, only the MCCM loses its signaling properties. Addition of exogenous TGF-β3 to target endocardial cells elicits an invasive migration only in cultures which have been activated *in vivo* by inductive interaction with the myocardium prior to treatment (Ramsdell and Markwald 1997). These results suggest that TGF-β3 may function to sustain and amplify cushion formation once induced by myocardial signals. In support of this, it was recently shown that both TGF-β Type II and Type III receptors are expressed in AV endothelial cells. Antibodies to both receptors inhibit endothelial to mesenchyme transition and mesenchymal cell migration (Brown, Boyer et al. 1996; Boyer, Erickson et al. 1999; Brown, Boyer et al. 1999).

As development proceeds, endothelial derived cushion cells colonize the cardiac jelly to form cellular outgrowths contributing to the formation of definitive heart chambers (Wunsch, Little et al. 1994). Cells targeted by the myocardial signal show a decrease in cell adhesion molecules (CAM) and an upregulation of serine and metalloproteinases, facilitating the loss of adhesion prior to matrix invasion. N-CAM is a cell adhesion molecule believed to be important in early cardiac cell-cell adhesion. Endothelial cells undergoing the transition to form mesenchyme, both *in vitro* and *in vivo*, show a decrease in N-CAM expression (Mjaatvedt and Markwald 1989). Alexander and coworkers demonstrated that the matrix metalloprotease, MMP-2, is expressed by the endocardium of the cushion tissues undergoing the transition to mesenchyme and by migrating mesenchymal cells. This suggests a role for MMP-2 in regulating cell motility

and matrix invasion (Alexander, Jackson et al. 1997). Additionally, elevated activity of urokinase, a serine protease, occurs with the onset of mesenchymal cell migration in the endocardial cushion tissues (McGuire 1990). Urokinase converts inactive plasminogen to plasmin, a broad-spectrum protease capable of activating procollagenase and degrading many components of the extracellular matrix (McGuire and Alexander 1993). Upregulation of genes involved in matrix interactions, such as fibronectin (Mjaatvedt, Lepera et al. 1987), fibulin-1 (Bouchey, Argraves et al. 1996), fibulin-2 (Miosge, Sasaki et al. 1998), proteoglycans (Funderburg and Markwald 1986; Little and Rongish 1995), tenascin (Crossin and Hoffman 1991) and hyaluronate synthase (Spicer, Augustine et al. 1996) also occurs. Hurle and coworkers found as endocardial cells delaminate to form the cushion tissue mesenchyme (HH stages 17-20), fine fibrillar tracts containing fibrillin-1, a component of 10-12 nm microfibrils, are detected in association with these cells (Hurle, Kitten et al. 1994). Similarly, emilin (elastic microfibril interface located protein) positive tracts extend through the cardiac jelly and are associated with the migrating cushion mesenchymal cells (Hurle, Kitten et al. 1994).

The now cellularized cushions expand by cell proliferation and send out extensions into atrial and ventricular regions to form septa. The atrial extension of the superior AV cushion (SAVC) comes in contact with the posterior wall of the atrium. Extracardiac mesenchyme (spina vestibuli) enters the atrium at this point of contact and has been shown to grow along the atrial extension adding tissue to the future atrial septum (Tasaka, Krug et al. 1996). The ventricular extension of the SAVC comes in contact with and fuses to a similar ventricular extension from the sinistro-ventral conal cushion (SVCC) (Figure 2). The fused mesenchymal tissue progressively migrates,

deepening the inner curvature fold of the ventricular myocardium (Mjaatvedt, Yamamura et al. 1999). The final morphogenetic step of the SAVC and SVCC is the muscularization of these tissues in a process known as myocardilization (Figure 2, beginning at HH stage 24). Nonproliferating myocardial cells from the inner curvature of the heart migrate into conal cushions and the SAVC allowing the posterior wall of the conus to fuse with the anterior wall of the right AV canal forming the mitroaortic continuity (Mjaatvedt, Yamamura et al. 1999).

The extracellular matrix plays a number of crucial roles during heart development. Initially, expansion of the endocardial cushions occurs by secretion of matrix proteins in the AV canal and OT. The matrix deposited allows the cardiac adheron, itself composed of ECM proteins, to signal the endocardial cells to begin their transformation to mesenchyme. Once transformed, the mesenchymal cells use, and add to, the matrix to migrate and populate the cushions eventually giving rise to the valves and septum needed to form a four chambered heart.

2.) Animal Models:

While the processes of cushion formation, epithelial-mesenchyme transition and myocardilization are occurring, the tubular heart has been looping to bring the developing septa into proper alignment to produce a normal four-chambered heart. Looping is under control of a variety of molecules, of which only a handful are known. The study of animals with *situs inversus* (reversal of left-right polarity) offers an opportunity to investigate the mechanism of looping.

Homozygosity of the *iv* (inversus viscerum) mutation in mice results in the randomization of the direction of heart looping (Brueckner, D'Eustachio et al. 1989).

Recently the gene was found and determined to encode a protein named left-right dynein (LRD) (Supp, Witte et al. 1997). Dyneins are a family of microtubule-based motors with axonemal dyneins producing ciliary and flagellar movement. In addition to *situs inversus*, there is a high frequency of cardiac defects. Approximately 20% of homozygous *iv/iv* mice exhibit heart defects, regardless of whether the heart is looping abnormally to the left or normally to the right (Brueckner, D'Eustachio et al. 1989). Interestingly, *situs inversus*, mucociliary dysfunction in the lungs and immotile sperm, are observed in the human autosomal recessive disorder known as Kartagener syndrome as a result of defects in the dynein arms of cilia (Fishman and Chien 1997).

The recessive *inv* mutation, in which all homozygous mice have *situs inversus*, resulted from the random insertion of a transgene (Yokoyama, Copeland et al. 1993). The gene disrupted by the transgene insertion was cloned, sequenced and published almost simultaneously by two separate groups (Mochizuki, Saijoh et al. 1998; Morgan, Turnpenny et al. 1998). The encoded protein contains 15-16 ankyrin-like repeats at the amino terminus believed to be involved in protein-protein interactions. A member of the TGF-β family called *lefty* is found expressed in the left half of gastrulating mouse embryos. In the *iv* and *inv* mouse mutants, *lefty* expression is inverted suggesting that *lefty* may function in determination of left-right (LR) symmetry downstream of *iv* and *inv* (Meno, Saijoh et al. 1996).

Other animal studies have helped in elucidating factors involved in looping. Levin and coworkers, working with chick embryos, described the asymmetric expression patterns of three genes involved in LR determination: *activin receptor IIa*, *Sonic hedgehog* (*Shh*), and *cNR-1* (Levin, Johnson et al. 1995). Reversing the sidedness of

either activin protein (normally on right) or *Shh* expression (normally on left) altered heart looping (Levin, Johnson et al. 1995). Repression or reversing the sidedness of *Shh* expression suggests that *Shh* activates expression of *cNR-1* (chicken *nodal related 1*) (Levin, Johnson et al. 1995).

Other mouse models have been important in determining genes involved in normal cardiac development. Mice homozygous for a null mutation of the MADS-box transcription factor *MEF2C* are embryonic lethal (Lin, Schwarz et al. 1997). In mutant embryos, the heart tube does not initiate a rightward looping and there is no morphological evidence for a future right ventricle. In addition, the left ventricle is severely hypoplastic, the trabeculae are poorly developed, the endocardial cells appear to be disorganized, and although the AV canal is present, endocardial cushions do not form (Lin, Schwarz et al. 1997).

Homozygous mice in which the homeobox gene *Nkx2-5* has been knocked out show growth retardation and die from cardiac insufficiency. A linear heart tube forms and begins to beat but fails to undergo correct looping. Additionally, null mice have diminished cardiac expression of the left ventricle marker, *eHand* (Biben and Harvey 1997). Once again, as with the *Mef2C* mouse, no endocardial cushions are formed and trabeculation is poor (Lyons, Parsons et al. 1995).

The bHLH transcription factors, dHAND and eHAND, show restricted expression during the looping process of cardiogenesis. dHAND is expressed only in the right atrium while eHAND expression is restricted to the conotruncus and left ventricle (Srivastava 1999). The MEF2 proteins have been shown to interact as cofactors with the MyoD family of bHLH proteins (Molkentin, Black et al. 1995). It has been speculated

that MEF2C acts as a cofactor with both dHAND and eHAND (Black and Olson 1999; Srivastava 1999). Such an interaction could account for the MEF2C null phenotype in which no right ventricle is formed (*dHand* is down regulated) and the left ventricle is hypoplastic (eHAND is present) (Lin, Schwarz et al. 1997). Conversely, *Nkx2-5* null mice show a decreased expression of eHAND while MEF2C expression is normal, suggesting a similar role for eHAND in the formation of the left ventricle.

The heart defect (*hdf*) mouse line arose from a recessive lethal insertional mutation on chromosome 13 (Yamamura, Zhang et al. 1997). The future outflow tract and right ventricle fail to develop normally and the endocardial cushion swellings in both the OT and AV canal are missing. Rescue experiments where AV endothelial cells are grown on three-dimensional collagen gels in the presence of myocardium from normal AV showed that the *hdf* endothelium is intrinsically competent to form cushion mesenchyme and that the defect is an extrinsic factor probably secreted by the myocardium. Fibronectin (FN), a matrix protein involved in cell adhesion and spreading, cell migration and cytoskeletal organization, is a major component of the endocardial cushions (Kitten, Markwald et al. 1987). FN is present in the matrix of homozygous *hdf* mice, but not in the same pattern as hemizygous *hdf* mice (Yamamura, Zhang et al. 1997), concluding that the mutant might be involved with the production, modification, or distribution of extracellular matrix molecules.

Mice that lack fibronectin die in early embryonic development with defects in mesodermally derived tissues. The notochord and somites are absent, the yolk sac, extraembryonic vasculature and amnion are defective, as well as the embryonic vessels and heart being abnormal and variable (George, Georges-Labouesse et al. 1993).

Correct looping of the heart is required to bring the septa into proper orientation and position to form a normal four-chambered heart. Some of the known defects of looping are caused by alterations of transcription factors, which affect a variety of downstream molecules. Defects in ciliary and flagellar movement, an ankyrin repeat containing protein and ECM proteins also have been shown to disrupt proper looping.

3.) **Human Genetics:**

As was mentioned earlier, CHD are the most common form of birth defect appearing in nearly one percent of newborn infants (Eisenberg and Markwald 1995). The most common of these are defects in atrial and ventricular septation. In spite of the high incidence of heart malformations, only 3 causal genes for congenital heart malformations have been identified. All three were characterized for rare disorders; the *jagged 1* (*JAG1*) gene in Alagille syndrome, *TBX5* in Holt-Oram syndrome, and *NKX2-5* in autosomal dominant ASD/atrioventricular conduction delay (Basson, Bachinsky et al. 1997; Li, Krantz et al. 1997; Li, Newbury-Ecob et al. 1997; Oda, Elkahloun et al. 1997; Schott, Benson et al. 1998).

Alagille syndrome is an autosomal dominant developmental disorder that affects structures in the liver, heart (tetralogy of Fallot), skeleton, eye, face, and kidneys (Li, Krantz et al. 1997). Identification of patients with cytogenetic deletions allowed mapping of the gene responsible to chromosome 20p12 and suggested that haploinsufficiency is one mechanism of causing Alagille syndrome (Oda, Elkahloun et al. 1997). Two groups simultaneously mapped the human *JAG1* gene to the Alagille critical region and linked mutations in the gene to the syndrome (Li, Krantz et al. 1997; Oda, Elkahloun et al.

1997). *JAG1* encodes a ligand for the developmentally important Notch receptor. The Notch signaling pathway controls the ability of non-terminally differentiated cells to respond to differentiation and proliferation signals (Artavanis-Tsakonas 1997).

Holt-Oram syndrome is characterized by upper limb malformations and cardiac septation defects (Basson, Bachinsky et al. 1997). Mutations in the *TBX5* gene, a member of the *Brachyury (T)* family of transcription factors, have been identified in five families and three sporadic cases of Holt-Oram syndrome (Basson, Bachinsky et al. 1997; Li, Newbury-Ecob et al. 1997). Li and coworkers investigated the expression of *TBX5* in human embryos between 26 and 52 days gestation. The highest cardiovascular expression was observed in the inflow of the heart tube (primitive atria and sinus venosus) at 26 days and later in the atrial wall, atrial septa, coronary sinus, and AV endocardial cushions and valves (33,41,48, and 52 days) (Li, Newbury-Ecob et al. 1997). This pattern of expression is consistent with the areas of the heart where structural defects of Holt-Oram syndrome (HOS) arise. Between 33 and 52 days, high levels of expression can be detected in developing forelimbs, trachea, lung and thoracic wall (Li, Newbury-Ecob et al. 1997). Structural defects are not often observed in the trachea and lung of HOS suggesting possible redundancy within the TBX gene family.

Recently, mutations in the gene encoding the human homeobox transcription factor NKX2-5 were found to cause nonsyndromic congenital heart disease and atrioventricular conduction abnormalities (Schott, Benson et al. 1998), the most prominent cardiovascular defect being atrial septal defects. Of the four families shown to have *NKX2-5* mutations, two had C-T transitions at nucleotide 642 that are likely to alter target-DNA binding. The other two mutations resulted in truncated proteins believed to

be unable to bind DNA, thus resulting in haploinsufficiency (Schott, Benson et al. 1998). *NKX2-5* is the homologue of the *Drosophila melanogaster* gene known as *tinman*, a homeobox transcription factor having an essential role in specifying heart muscle progenitors in nascent mesoderm (Schott, Benson et al. 1998).

More commonly, CHD occur as part of a malformation syndrome caused by chromosomal aberrations, resulting in abnormal dosage of one or more genes. Trisomy 21 (Down syndrome) accounts for the majority of congenital heart defects associated with chromosomal abnormalities (Kramer, Majewski et al. 1987). Liveborn children with Down syndrome (DS) have a 50-fold increased incidence of congenital heart defects (Klewer, Krob et al. 1998). When specifically looking at endocardial cushion defects (ECD), Down syndrome constitutes 78% of the syndromic ECD and 59.5% of all ECD (Carmi, Boughman et al. 1992). The remaining cases of ECD are isolated (24%) or associated with other syndromes or chromosomal abnormalities (Carmi, Boughman et al. 1992). The DS phenotype and associated ECD are most likely caused by overexpression of a number of genes on chromosome 21 due to additional gene copies. Phenotypic features of DS have been mapped to distinct regions of chromosome 21 through the use of families with partial trisomy.

In trisomy 21, the critical region for endocardial cushion defects has been narrowed to a 9 Mb span of DNA localized to chromosome 21q22.2-21q22.3 (Payne, Johnson et al. 1995). Although the gene(s) responsible for cushion defects in DS have not been identified, the collagen type VI, α-1 and α-2 genes, map to the critical region and are excellent candidate genes. The collagen genes are coordinately regulated and expressed in the human fetal heart (Duff, Williamson et al. 1990). The collagen protein

has been shown to localize within the embryonic AV valves (Hurle, Kitten et al. 1994; Kitten, Kolker et al. 1996) and be involved in proliferation and migration of cells (Perris, Kuo et al. 1993; Pfaff, Aumailley et al. 1993; Atkinson, Ruhl et al. 1996). Further, Kitten and coworkers demonstrated that addition of antibodies to type VI collagen inhibited the attachment and migration of the transformed mesenchyme cells (Kitten, Kolker et al. 1996).

Chromosomal aberrations resulting from deletions may also give rise to cardiac defects. Congenital conotruncal defects observed in DiGeorge syndrome arise from haploinsufficiency of one or more genes present on chromosome 22q11 (Farrell, Stadt et al. 1999). Microdeletions of chromosome 22q11 are the most common genetic defects associated with cardiac and craniofacial anomalies in humans and cause DiGeorge, velo-cardio-facial (VCFS) and conotruncal anomaly face syndromes (CAFS) (Yamagishi, Garg et al. 1999). Deletions in this region are believed to affect cardiac neural crest migration and function. Recently, Yamagishi and coworkers suggested the human *UFD1L* gene, which encodes a protein involved in degradation of ubiquitinated proteins, can contribute to many of the congenital heart and craniofacial defects seen in 22q11 microdeletion syndrome (Yamagishi, Garg et al. 1999). The gene was deleted in all 182 patients studied with 22q11 deletion. In addition, they screened a number of individuals with cardiac and craniofacial defects who did not have detectable deletions and found one individual with a de novo deletion of exons 1 to 3 of *UFD1L* (Yamagishi, Garg et al. 1999).

There is substantial intra- and inter-familial variability in the phenotype associated with the 2 Mb 22q11 deletion suggesting other genes are likely involved. The

human *HIRA* gene may contribute some to the variable phenotype. The gene was named after the yeast histone regulatory genes, which act as repressors of histone gene transcription (Schiaffino, Dallapiccola et al. 1999). It has also been shown that HIRA orthologues, in chick and mouse, are expressed in neural crest cells and neural crest derived tissues (Farrell, Stadt et al. 1999). Farrell and coworkers using antisense oligonucleotides to attenuate *cHIRA* in chick cardiac neural crest, ex-ovo, followed by orthotopic backtransplantation to untreated embryos, found an increased incidence of persistent truncus arteriosus, a characteristic of DiGeorge syndrome (Farrell, Stadt et al. 1999). However, they did not observe any affect in the repatterning aortic arch arteries, the ventricular function, or the alignment of the outflow tract. Homozygous inactivation of the *Hira* gene in mice results in death about embryonic day 10, before cardiac neural crest migration and outflow septation (Scambler, Roberts et al. 1998).

Individuals with deletion of the distal region of chromosome 8p have CHD (typically ECD) in addition to microcephaly, intrauterine growth retardation, mental retardation and a characteristic hyperactive, impulsive behavior (Devriendt, Matthijs et al. 1999). The human gene encoding the GATA4 transcription factor maps to chromosome 8p23.1-p22 (Huang, Heng et al. 1996) and is known to be important in ventral morphogenesis and heart formation (Kuo, Morrisey et al. 1997; Molkentin, Lin et al. 1997). Devriendt and coworkers found that *GATA4* was deleted in all their patients except for one individual in which no heart defect was observed, making it an excellent candidate gene for CHD (Devriendt, Matthijs et al. 1999).

Although the genetic basis for some forms of CHD have been determined, many more remain unknown. In addition to the single gene defects mentioned earlier, a

number of families have shown endocardial cushion defects (ECD), also known as AV canal defects (AVCD), being transmitted in an autosomal dominant fashion with incomplete penetrance (O'Nuallain, Hall et al. 1977; Emanuel, Somerville et al. 1983; Wilson, Curtis et al. 1993; Cousineau, Lauer et al. 1994; Gennarelli, Novelli et al. 1994; Amati, Mari et al. 1995; Johnson, Payne et al. 1995; Payne, Johnson et al. 1995; Burn, Brennan et al. 1998). Under the heading of ECD/AVCD lie the malformations known as atrial septal defects, ventricular septal defects, complete atrioventricular canal defects, and common atrium. The spectrum of AVCDs is attributed to abnormal extracellular matrix during development (Lin, Herring et al. 1999). Mapping with some families has shown exclusion of linkage with chromosome 21 and 8 (Wilson, Curtis et al. 1993; Cousineau, Lauer et al. 1994; Gennarelli, Novelli et al. 1994; Amati, Mari et al. 1995), suggesting alterations in genes elsewhere in the genome are responsible.

4.) 3p- Syndrome:

The first patient with monosomy 3p25-pter, due to a deletion of the distal part of chromosome 3 (3p- syndrome), was presented by Verjall and De Nef in 1978 (Verjall and Nef 1978). Since then, only 22 additional cases have been presented in the last two decades making this syndrome very rare. This is interesting, since Aula and von Koskull found that chromosome 3p2 is one of the most common sites for spontaneous chromosome breakage in lymphocyte cultures, accounting for 13% of all observed breaks (Aula and Koskull 1976). However, 3p- has not been observed in studies of spontaneous abortions (Hassold, Chen et al. 1980; Kajii, Ferrier et al. 1980; Olson and Magenis 1988). Zygotes with 3p2-pter deletions may be aborted before the pregnancy is recognized

clinically (Merrild, Berggreen et al. 1981). In the few individuals who survive to birth, it is possible that the position of the break allows development to continue.

Characteristic features of 3p- syndrome include low birth weight, pre- and postnatal growth delay, psychomotor and mental retardation, microcephaly, ptosis, low set malformed ears, micrognathia, telecanthus, long philtrum, and hypotonia. More variable features include CHD, postaxial polydactyly, cleft palate, renal anomalies, gastrointestinal anomalies, rocker bottom feet, seizures, triangular face, preauricular pits, and hearing impairment (Verjall and Nef 1978; Merrild, Berggreen et al. 1981; Sagredo, Castilla et al. 1981; Higginbottom, Mascarello et al. 1982; Beneck, Suhrland et al. 1984; Witt, Biedermann et al. 1985; Reifen, Gale et al. 1986; Tolmie, Batstone et al. 1986; Ramer, Ladda et al. 1989; Tazelaar, Roberson et al. 1991; Mowrey, Chorney et al. 1993; Phipps, Latif et al. 1994; Drumheller, McGillivray et al. 1996). With the exception of one incidence in which a mother and son had deletion of 3p25-pter (Tazelaar, Roberson et al. 1991), all cases have arisen *de novo*. The mortality rate among reported cases is 25%, principally related to cardiac malformation, with all deaths occurring in infancy (Ramer, Ladda et al. 1989).

The cytogenetic breakpoint of the deletion associated with 3p- syndrome has been identified as 3p25. It is thought that the extent of the deletion may correspond with the severity of the syndrome. Using molecular genetic analysis, Phipps and co-workers investigated five cases of 3p- syndrome to determine a relationship between breakpoints and clinical phenotype (Phipps, Latif et al. 1994). It was determined that loss of sequence telomeric to D3S1317 was required for expression of the characteristic 3p- syndrome phenotype. An individual studied by Mowrey and co-workers helped define

the distal end of the critical region responsible for expression of characteristic features of 3p- syndrome when deleted. The individual was determined to have an interstitial deletion of chromosome 3p25-26 with the proximal breakpoint near the Von Hipple-Lindau (VHL) locus and the distal breakpoint near D3S17 (Mowrey, Chorney et al. 1993). The smallest critical region (21 cM) for expression of characteristic features of 3p- syndrome therefore must lie between D3S1317 and D3S17 (figure 3).

Other correlations between breakpoints and phenotype have been made. Molecular genetic analysis of five 3p- syndrome patients, three with cardiac septal defects, demonstrated that the patients with cardiac defects had more proximal deletions than those without (Phipps, Latif et al. 1994). Detailed molecular analysis of the chromosome breakpoints for each of the patients showed that the presence of CHD correlates with deletion of the interval between markers D3S1250 and D3S18 (4 cM). It was concluded that a gene involved in normal cardiac development resides in this interval and that deletion or disruption of that gene results in cardiac septal defects. At the time, three genes had been isolated from this region including the plasma membrane calcium transporting ATPase isoform 2 gene (PMCA2 or ATP2B2) (Latif, Duh et al. 1993), the VHL disease gene, and a cDNA (g6) of unknown function (Latif, Tory et al. 1993). Large germline deletions containing both the VHL and g6 genes were found in VHL disease patients who did not have congenital heart disease (Latif, Tory et al. 1993; Phipps, Latif et al. 1994) therefore eliminating them as candidates for CHD. Using fluorescent *in situ* hybridization (FISH) and polymorphic microsatellite analyses, Drumheller and co-workers refined the location of the putative cardiac development gene to a much smaller region bordered by D3S1585 and D3S1317 (2 cM) (Figure 3)

(Drumheller, McGillivray et al. 1996). In addition to PMCA2, the human homologue of the yeast Sec13 gene involved in vesicle formation is contained within this region. Detailed mapping by Green and coworkers refined the critical region to an even smaller interval (1000 kb) and excluded the candidate genes PMCA2 and fibulin-2 (Green, Latif et al. 1998).

The genetic factors involved in the complex process of cardiac development are numerous and only beginning to be understood. Identification and functional characterization of the genes expressed during cardiac development, and their protein products, are key to our understanding of normal heart development and the many routes to congenital heart defects.

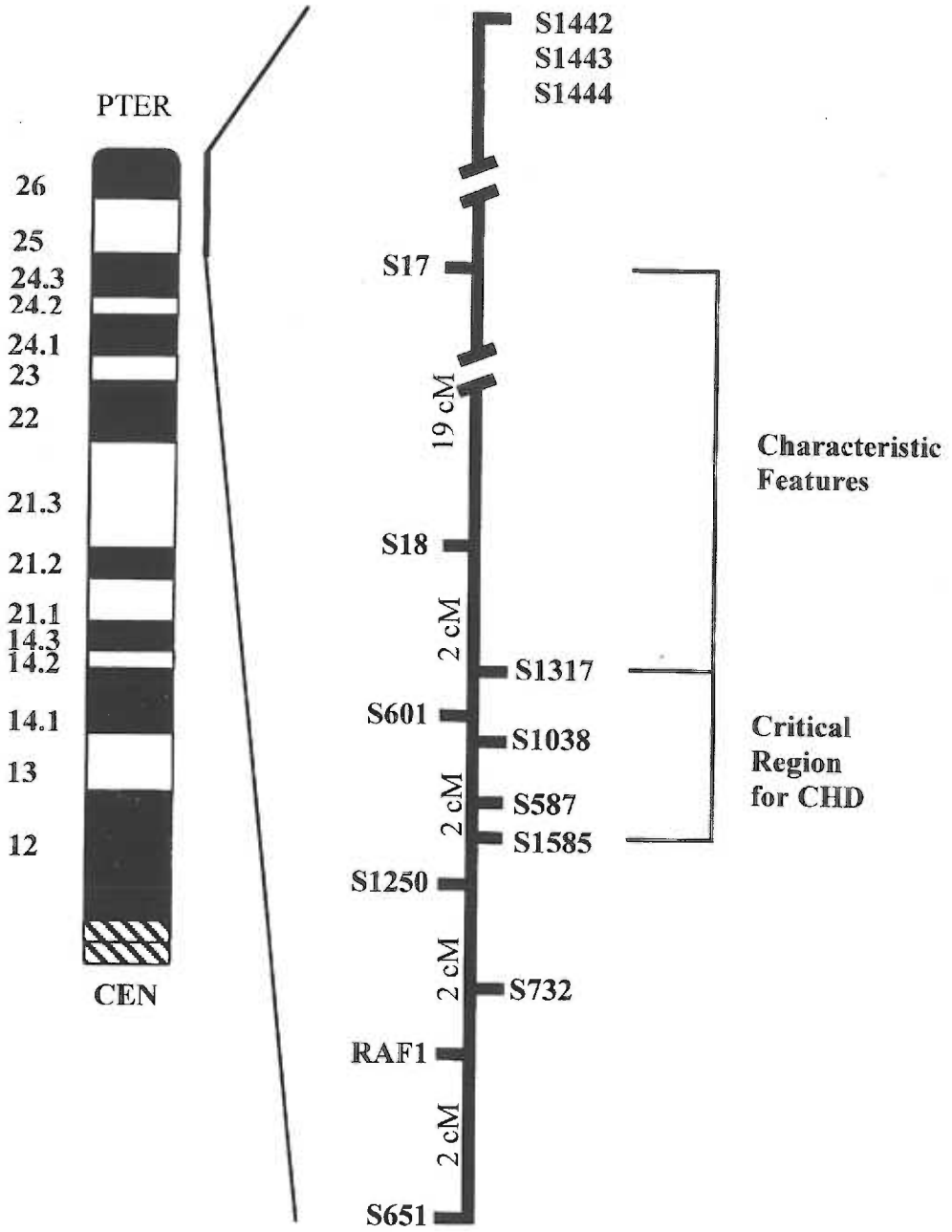


Figure 3. Diagram delineating critical chromosomal regions for congenital heart defects and characteristic features of 3p- syndrome.

5.) Hypothesis:

A screen of expressed sequence tag (EST) clones mapped to human chromosome 3p24.2-25 revealed a cDNA with sequence similarities to the fibrillins and other extracellular matrix proteins (Timmers, Whitney et al. 1996). This cDNA clone was further characterized based on the idea that ECM proteins play an important role in cardiovascular development. We have partially characterized the gene and its protein product, which we have named cirrin. Due to its chromosomal location, expression pattern, predicted protein structure, and potential association with ECM (described in this work), we hypothesize that cirrin is involved in normal cardiac development and that mutation or deletion of cirrin results in endocardial cushion defects.

Chapter 2

Materials and Methods

Whole-mount *in situ* hybridization

1.) 1X PBS (DEPC treated)

2.) 1X PBST (1X PBS + 0.1% Tween-20) (DEPC treated)

3.) 1X TBST (1X TBS + 0.1% Tween-20) (DEPC treated)

4.) 4% paraformaldehyde in PBS (make fresh every use)

5.) 25%, 50%, 75% Methanol/PBST; 100% Methanol

6.) Proteinase K (10 µg/ml in PBST)

7.) Hybridization mix:

	Final Concentrations
- formamide	50%
- SSC	1.3X SSC
- EDTA (pH 8.0)	5mM
- Yeast RNA	50 µg/ml
- Tween-20	0.2%
- CHAPS	0.5%
- Heparin	100 µg/ml
- DEPC dH_2O	XX ml
Total	50 ml

8.) Pre-block solution:

- 10% sheep serum (heat-inactivated at 70 °C x 10 min.)
- 100 µl BSA
- 8.9 ml TBST
- keep serum and BSA cold

9.) NTMT solution (make fresh with every use)

	Final Concentrations
- NaCl	100 mM
- Tris·Cl (pH 9.5)	100 mM
- $MgCl_2$	50 mM
- Tween-20	0.1 %
- dH_2O	XX
Total	50 ml

Embryo preparation:

Fertilized chicken embryos were incubated at 37 °C. Eggs were rotated one half turn every 12 hours. Embryos were harvested at 66, 72, and 90 hour time points. Embryos were resected and placed in DEPC treated PBS. Extraembryonic membranes were removed and embryos fixed in 4% paraformaldehyde at room temperature for two hours. Embryos were dehydrated for one hour sequentially with 25%, 50%, 75%, and 100% methanol/PBS then stored at -20 °C. Prior to hybridization, embryos were rehydrated for one hour with 25%, 50%, 75% and 100% PBS/methanol.

Probe preparation (Day 1):

Whitehead Institute clone 11041, which contains the complete cirrin gene, was digested with BamHI and AccI restriction enzymes. A fragment corresponding to cirrin sequence –321 to 732 was gel purified and subcloned into pGEM-4Z (Promega). The clone was linearized with BamHI or AccI for probe production. SP6 RNA polymerase, in conjunction with the AccI linearized plasmid as template, was used to produce a sense RNA probe (negative control). T7 polymerase, in conjunction with the BamHI linearized plasmid as template, was used to produce an antisense RNA probe (Figure 4).

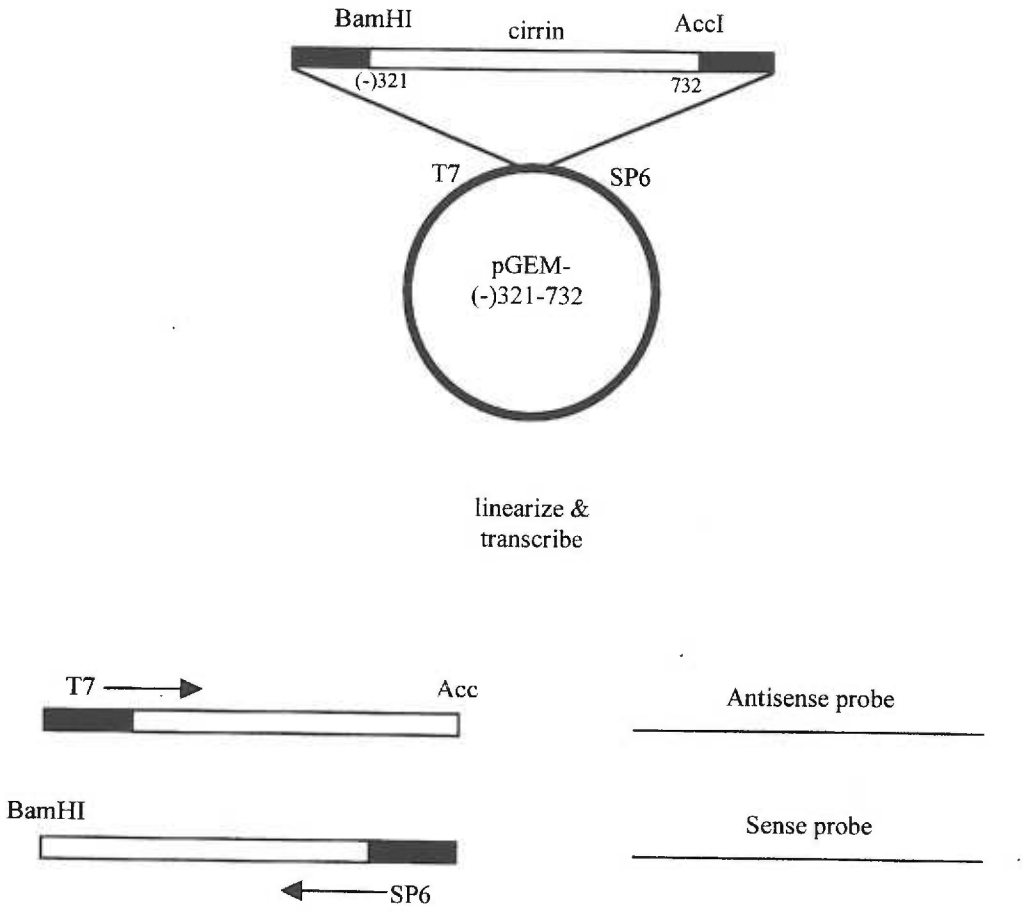


Figure 4. Diagram of construct used to produce sense (- control) and antisense (+) RNA probes for whole-mount *in situ* hybridizations. A 1053 base probe is created using cirrin cDNA sequence, from position (-) 321 to position 732, as template.

The following reactions were run to produce labeled probes:

x µl linearized plasmid (equivalent to 1 µg of insert DNA)
1 µl 10x nucleotide mix (with digoxigenin-UTP)
2 µl 5x transcription buffer
1 µl DTT (100 mM for T7/ 10 mM for SP6)
1 µl RNAsin
2 µl T7 or SP6 RNA polymerase (10 U/µl)
x µl DEPC dH_2O
10 µl total

- incubated at 37 °C for 1 hr.
- added 2 µl DNAse I (20 U/µl)
- incubated at 37 °C for 20 min.
- checked probes by electrophoresis on 0.8% agarose gel in TBE buffer
- aliquoted probes and stored at –20 °C

Hybridization:

- rehydrated embryos were incubated 2 x 30 min. in PBS
- 1 ml of 10 µg/ml proteinase K in PBST was added to each embryo
- embryos were incubated at 37 °C for 7-20 min. depending on size and stage of embryo
- to quench reaction, proteinase K was removed by pipetting and embryos washed 3 x 1 min. in PBST
- embryos were fixed in 4% paraformaldehyde at room temperature for 20 min. (rocking)
- embryos were rinsed 1 min. in PBST
- embryos were washed 10 min. in PBST

- embryos were placed in a 1:1 mix of PBS/hybridization solution at 70 °C x 30 min. (rocking)
- placed embryos in 100% straight hybridization solution at 70 °C x 1 hr. (rocking)
- put fresh hybridization solution on embryos
- added 500 ng of appropriate Dig-labeled probe to each embryo
- incubated at 70 °C x overnight

Washes and antibody hybridization (Day 2):

---Embryos should be transluscent after overnight hybridization---

- washed 2 x 1 min. with 70 °C hybridization solution
- washed 2 x 30 min. in 70 °C hybridization solution
- washed 30 min. in 1:1 mix of hybridization solution/TBST at 70 °C
- washed 3 x 1 min. in TBST at room temperature
- incubated 20 min. in 100 µg/ml RNAse A in TBST at 37 °C
- antibody pre-adsorption was begun now (described in next section)
- washed 2 x 1 min. with TBST at room temperature to remove RNAse A
- placed embryos in pre-block solution at 4 °C for 2 hr. (rocking)
- removed pre-block solution and replaced with pre-adsorbed antibody
- incubated overnight at 4 °C while rocking

Antibody pre-adsorption (anti-digoxigenin):

Mixed- 250 µl TBST
2.5 µl sheep serum (heat inactivated)
2.5 µl 10% BSA
10 mg of chick embryo powder (dehydrated ground up embryos)

- incubated at 70 °C x 20 min.
- vortexed 5 min at room temp
- iced 5 min.
- added 2 µl of Anti-Digoxigenin-AP Fab fragments and incubated at 4 °C x 2 hr. (rocking)
- centrifuged at 4 °C for 10 min. at 14,000 RPM
- removed supernatant and added to:

  9.8 ml TBST
  100 µl BSA (10%)
  100 µl Sheep serum (heat inactivated)

- mixed and kept on ice

Post-antibody washes and immunohistochemistry (Day 3):

- rinsed 3 x 1 min. in TBST
- washed 3 x 1 hr. in TBST
- rinsed 2 x 1 min. in NTMT (made fresh)
- washed 2 x 10 min. in NTMT
- incubated with NTMT buffer containing 4.5 µl/ml nitroblue tetrazolium (NBT) and 3.5 µl/ml bromochloroindolyl phosphate (BCIP) until color appeared (incubated in dark at room temperature)
- quenched reaction by rinsing 3 x with PBST
- fixed color by incubating in 4% paraformaldehyde in PBS overnight
- washed embryos 3 x 1 min. in PBS and photographed with Ektachrome 320T Tungsten film (Kodak)

Materials:

	Company	Cat. #
1.) DIG RNA labeling mix	Boehringer Mannheim	1277 073
2.) SP6 RNA Polymerase	GibcoBRL	18018-010
3.) T7 RNA Polymerase	GibcoBRL	18033-019
4.) Yeast RNA	GibcoBRL	15401-011
5.) CHAPS	Sigma	C-3023
6.) Heparin	Sigma	H-3393
7.) Sheep Serum	Sigma	
8.) Paraformaldehyde	Sigma	P-6148
9.) BSA	Sigma	A2153
10.) Anti-Digoxigenin-AP Fab fragments	Boehringer Mannheim	1093 274
11.) DNAse I	GibcoBRL	18047-019

Southern blot hybridization

A multiple species Southern blot (Zoo blot, Clontech) was probed with a cirrin cDNA probe covering exons 1-5. The probe was labeled with 50 μCi [α^{32}P] dCTP, 3000 Ci/mmol, using a random primed DNA labeling kit (Boehringer Mannheim) as per protocol. The percent incorporation of dCTP into the probe was determined to be 70.0 % by TCA precipitation. Unincorporated nucleotides were removed by centrifugation through ProbeQuant™ G-50 micro columns (Pharmacia Biotech) as per protocol. Hybridization was carried out at 60 °C for one hour in Expresshyb hybridization solution (Clontech) as per protocol. The blot was washed with four changes of 2X SSC/0.5% SDS at room temperature for 40 minutes (4 x 10 min. washes). Further washes were carried out at 65 °C with three changes of 0.1X SSC/0.1% SDS for 60 minutes (3 x 20 min. washes). The blot was exposed to DuPont-NEN Reflection film for 15 hours at –80 °C and then developed.

Northern blot hybridization

Poly-A fetal and adult multiple tissue northerns (Clontech) were probed with the entire coding sequence of cirrin including 321 bases of 5'-UTR and 357 bases of 3'-UTR. The probe was labeled using the Gene Images labeling module (Amersham) as per protocol. Hybridization was carried out at 65 °C for 12 hours. Blots were washed at 65 °C for 15 minutes in both 1X SSC/0.1% SDS followed by 0.1X SSC/0.1% SDS. Cirrin transcripts were detected using the Gene Images CDP-Star detection module (Amersham) as per protocol. Blots were exposed to Hyperfilm ECL (Amersham) for 45 minutes and

then developed. Blots were stripped to remove the probe and then reprobed as above with a labeled β-actin probe to check for equality of sample loading.

RNA Master Blot

An RNA master blot (Clontech) was probed with a cDNA probe covering exons 1-5 of cirrin. The probe was labeled with 50 μCi [α^{32}P] dCTP, 3000 Ci/mmol, using a random primed DNA labeling kit (Boehringer Mannheim) as per protocol. The percent incorporation of dCTP into the probe was determined to be 54% by TCA precipitation. Unincorporated nucleotides were removed by centrifugation through ProbeQuant™ G-50 micro columns (Pharmacia Biotech) as per protocol. Hybridization was carried out at 68 °C for one hour in Expresshyb hybridization solution (Clontech) as per protocol. The blot was washed with four changes of 2X SSC/0.5% SDS at room temperature for 40 minutes (4 x 10 min. washes). Further washes were carried out at 65 ° with two changes of 0.1X SSC/0.1% SDS for 40 minutes (2 x 20 min. washes). The blot was exposed to DuPont-NEN Reflection film for 240 hours at –80 °C and developed.

cDNA sequence analysis

Clone 11041 from the Whitehead Institute was isolated from an overnight culture of DH5α cells using a plasmid mini-prep kit from Qiagen as per protocol. The clone was sequenced with an ABI 377 automated fluorescence sequencer using the T3, T7, 2-seq, 3-seq, 4-seq, 5-seq, 10-seq, and MG311 primers (Appendix). cDNA sequence was confirmed from multiple reactions on total genomic DNA.

Determination of intron-exon structure

Amplification of the carboxy terminal half of cirrin from BAC 172I17 using primers 2T7 and 2SP6 was accomplished with the following reaction:

---note: all concentrations are initial

20.75 µl dH_2O
15.50 µl genomic gelatin buffer *
1.00 µl BAC 172I17 (45 ng)
1.00 µl 2T7 (25 µM)
1.00 µl 2SP6 (25 µM)
10.00 µl dNTP's (10 mM)
0.50 µl DMSO
0.25 µl Taq polymerase (5 U/µl)
50 µl total

94 °C/5' - [94 °C/1' - 60 °C/1' - 72 °C/5'] X 35 cycles – 72 °C/10'

*genomic gelatin buffer: (50 µl reaction)
3 mM $MgCl_2$
10 mM Tris·Cl (pH 8.4)
50 mM KCl
0.1 µg gelatin

Introns 5-9 were found by sequencing the 2T7/2SP6 PCR product using an ABI Prism 377 automated fluorescence DNA sequencer with 6-seq, 7-seq, 10-seq, 11-seq, 14-seq, MG311, and 2T7 primers (Appendix).

Amplification using the Expand™ Long Template PCR System (Boehringer Manheim) on BAC 172I17 yielded an approximately 3200 base product (371 bases in cDNA) using primers 12-seq and 1-2R with the following reaction:

14.30 µl dH_2O
5.00 µl Expand buffer 1 (17.5 mM $MgCl_2$)
2.00 µl BAC 172I17 (90 ng)
0.60 µl 12-seq (25 µM)
0.60 µl 1-2R (25 µM)
1.75 µl dNTP's (10mM)
0.75 µl Taq/Pwo polymerase mix (3.5 U/µl)
25.0 µl total

94 °C/2' – [94 °C/20" – 62 °C/30" – 68 °C/10'] X 30 cycles – 68 °C/15'

The 12-seq/1-2R PCR product was cloned using the pMOS*Blue* blunt ended cloning kit (Amersham). Introns 1 and 2 were found by sequencing the constructed clone using primers U19, 2-seq, and T7 (Appendix).

To find the remaining introns, BAC 172I17 was digested with Pst I restriction enzyme and cloned into the pGEM-4Z vector (Promega) using standard methods of ligation and transformation. Colony lifts onto NEN Life Science Colony/Plaque Screen hybridization Transfer Membrane were probed with a cDNA PCR product spanning the region of unknown intron structure (2-1F and 13-seq primers). The PCR product was labeled with 50 µCi [α^{32}P] dCTP, 3000 Ci/mmol, using a random primed DNA labeling kit (Boehringer Mannheim) as per protocol. Introns 3 and 4 were found by sequencing positive clones with 15-seq, 19-seq and 20-seq primers (Appendix 1).

Preparation of polyclonal antibody 2153

A peptide spanning cirrin amino acids 81-98 was synthesized and sequenced to determine if correct. For subcutaneous injection into a New Zealand white rabbit, 3 mg of peptide was coupled through its terminal cysteine to a carrier protein (keyhole limpet

hemocyanin) using the Imject® Maleimide Activated Carrier Proteins from Pierce as per protocol. Coupled peptide (200 µl) was mixed with an equal amount of TiterMax® Research Adjuvant (CytRx Corporation) as per protocol and injected. Ten milliliters of pre-immune serum was collected prior to injection and stored at –20 °C. Antibody titer was checked by ELISA assay every two weeks when immune serum was collected using the peptide as antigen. The rabbit was boosted if the titer was low. Immune, as well as preimmune, serum was purified through a protein G column using IgG Binding Buffer and IgG Elution Buffer from Pierce. A peptide column was prepared using the SulfoLink® Kit from Pierce as per protocol. Immune serum was purified over the affinity column.

Preparation of polyclonal antibody 1851

A peptide spanning cirrin amino acids 61-95 was synthesized and sequenced to determine if correct. For subcutaneous injection into a New Zealand white rabbit, 0.50 mg of peptide was suspended in 300 µl PBS and then mixed with an equal amount of TiterMax research adjuvant (CytRx Corporation) as per TiterMax protocol. Preimmune and immune serum was collected as described above and stored frozen. Antibody response was followed by ELISA using synthesized peptide (amino acids 61-95) as an antigen. Immune, as well as preimmune, serum was purified through a protein G column using IgG Binding Buffer and IgG Elution Buffer from Pierce. Immune serum was purified over a HiTrap affinity column (Amersham) coupled to synthetic peptide 61-95 as per protocol.

Baculovirus expression of cirrin

For recombinant baculovirus expression of cirrin, the coding sequence of cirrin was cloned into pFastBac Hta donor plasmid (GibcoBRL). Production of a recombinant cirrin viral stock was accomplished following the Bac-To-Bac Baculovirus Expression Systems protocol (GibcoBRL). A viral titer was determined using the BacPAK™ Rapid Titer Kit (Clontech) as per protocol. A 50 ml culture of Sf9 cells at 2 x 10^6 cells/ml were infected at a multiplicity of infection (MOI) equal to 0.2 and 2.0 and grown in the presence of E64, leupeptin, and pepstatin protease inhibitors at 27 °C and 130 RPM. At 48, 72, 96, 120, and 144 hours post-infection, 10 ml of culture was collected. The cells were spun down at 500 x g for 10 min. and the supernatant saved at 4 °C. To isolate protein from the cell, the cell pellet was resuspended in 5 volumes lysis buffer (50 mM Tris-HCl (pH 8.5), 1% Nonidet P-40, 10 mM β-mercaptoethanol, 1 μM phenylmethylsulfonyl fluoride(PMSF)) per gram of cells and processed as per Bac-To-Bac Expression system protocol. pFastBac Hta donor plasmid without insert was processed as a negative control.

Isolation of cirrin from cartilage extract

For native fetal bovine cirrin, the cartilagenous ends of the long bones of fetal calves were collected, frozen, pulverized and extracted sequentially with 50 mM Tris-HCl, pH 7.5 containing first 0.2 M NaCl, then 1.0 M NaCl, then 1 M NaCl and 20 mM EDTA, then 6 M urea and finally 1.0% SDS. Each extract contained 1 mM of the protease inhibitors 4-(2-Aminoethyl)benzenesulfonyl fluoride (AEBSF), N-

Ethylmaleimide (NEM) and benzamidine. Analysis of each fraction by western blot analysis showed that cirrin is extracted in the first low salt wash (not shown).

In vitro isolation of cirrin

To isolate cirrin secreted into tissue culture medium, human normal skin fibroblasts were grown to 75% confluence in α-MEM/15% FBS/0.1 % gentamicin at 37 °C and 5% CO_2. Medium was removed and cells washed two times with PBS. Cells were then incubated in serum free medium for 24 hrs. Three separate aliquots (1.0, 0.5, and 0.25 ml) of the conditioned serum free medium were TCA precipitated. One tenth volume of 15% deoxycholate was added to the medium, mixed, and allowed to sit at room temperature for 10 min. One tenth volume of ice cold 72% TCA was added and protein pelleted by centrifugation for 10 min. at 3000 x g. Supernatant was removed and pellet resuspended in 30 µl of 100 mM NaOH.

Western blot analysis

For western blot analysis of baculovirus and cartilage extract, SDS-PAGE sample buffer was added to each sample and 10 µl of the cartilage fraction and 10 µl of each insect cell protein extract were separated on 8.5% SDS-PAGE and transferred to PROTRAN nitrocellulose membrane (Schleicher and Schuell). To check for antibody specificity, primary antibody diluted 1:100 was preincubated with 40 mg/ml antigen at 4 °C for 2 hrs. prior to probing a duplicate blot. The membranes were blocked in PBS with 5% non-fat dry milk and immunoprobed with a 1:100 dilution of affinity purified

antibody 1851 or the preblocked antibody. Blots were developed using the ECL western blotting system (Amersham) as per protocol.

For western blot analysis of tissue culture medium, SDS-PAGE sample buffer was added to 10 µl of each precipitated sample, separated by 8.5% SDS-PAGE and transferred to PROTRAN nitrocellulose membrane (Schleicher and Schuell). The membrane was blocked in PBS with 5% non-fat dry milk and immunoprobed with affinity purified antibody 1851 (1:100 dilution). After incubation with goat anti-rabbit IgG-AP conjugated antibody (1:3000 dilution) (BIO-RAD), the blot was developed using BCIP/NBT detection reagents (Promega).

Fluorescent *in situ* hybridization (FISH)

A genomic probe, specific for the cirrin gene, was engineered for FISH analysis. BAC 172I17 was digested with EcoRI restriction enzyme and fragments cloned into pGEM-4Z vector (Promega). Plasmid isolated from individual clones was digested with EcoRI, run on a gel and transferred to Hybond-N+ membrane (Amersham) for Southern blot analysis. The blot was probed with a full length cirrin cDNA probe. A single clone (clone-21) with an insert of approximately 11Kb was identified. DNA from a bacterial overnight culture containing clone-21 was isolated using the QIAprep Spin Miniprep Kit (Qiagen) as per protocol for use as a probe in FISH analysis. Positive PCR results using clone-21 as template with primers spanning exons 1 (8F/8R) and 10 (7F/7R) suggest that it encompasses the entire coding region of cirrin (exons 1-10). Sequencing of clone-21 with primer 8F indicated that exons 1-2 and intron 1 were present within the clone.

Three well-characterized cell lines (GM07873, GM10922, and GM10985) with terminal deletions of the short arm of chromosome 3 at band p25 (3p-) were obtained from the NIGMS Human Mutant Cell Repository (Coriell Institute). Lymphoblast cell lines (GM10922 and GM10985) were maintained in RPMI-1640 medium with 15% heat-inactivated FBS and 0.1% gentamicin at 37 °C and 5% CO_2. Fibroblast cell line GM07873 was maintained in α-MEM with 15% FBS and 0.1% gentamicin at 37 °C and 5% CO_2.

FISH analysis of the cell lines, with the cirrin probe, was performed by Dr. Susan Olson and Carol Reifsteck. 1320 ng of the clone-21 cirrin probe was nick translated with digoxygenin-11-dUTP (Boehringer-Mannheim). Metaphase spreads from each cell line were hybridized with 200 ng of cirrin probe with blocking agent and a chromosome 3 alpha satellite probe labeled with biotin (Oncor), as an identifier. The probes were detected with Oncor reagents, anti-digoxygenin-rhodamine/anti-digoxygenin-FITC for the cirrin probe and avidin-FITC for the alpha satellite probe. Since the cirrin probe target was small (11 kb), localization was confirmed using two different colors for detection (rhodamine, red and FITC, yellow) in separate FISH experiments.

Patient southern blot analysis

Ten micrograms each of genomic DNA from GM10922, GM10985, and two normal controls were digested with 50 U of either PstI or EcoRI for 2 hrs. at the appropriate temperature. The digested DNA was electrophoresed on a 0.8% agarose gel in TBE gel at 5 V/cm. The DNA was transferred to a Hybond-N+ membrane (Amersham) for Southern blot analysis as per standard protocols. The blot was probed

with the entire coding sequence of cirrin including 321 bases of 5'-UTR and 357 bases of 3'-UTR. The probe was labeled using the Gene Images labeling module (Amersham) as per protocol. Hybridization was carried out at 60 °C for 18 hours. The blot was washed at 60 °C for 15 minutes in both 1X SSC/0.1% SDS followed by 0.5X SSC/0.1% SDS. Probe binding was detected using the Gene Images CDP-Star detection module (Amersham) as per protocol. The blot was exposed to Hyperfilm ECL (Amersham) for 45 minutes and then developed.

Homology and domain searches

Homology to known protein domains was determined using the basic BLAST 2.0 algorithm located at the National Center for Biotechnology Information (NCBI) homepage (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST). The search was performed using the entire coding sequence as well as each exon individually or tandem pairs using the basic program parameters. Comparisons were done against DNA and protein databases. The amino acid sequence was also analyzed through the Tmpred algorithm for potential transmembrane domains (http://www.ch.embnet.org/software/TMPRED_form.html), and the PSORT II server for prediction of protein sorting signals (http://psort.nibb.ac.jp:8800/form.html). Optimal alignment of cirrin, HT and F09E8 protein sequences was done using the multiple sequence alignment program MSA version 2.1 (http://www.ibc.wustl.edu/ibc/cgi-bin/msa.cgi).

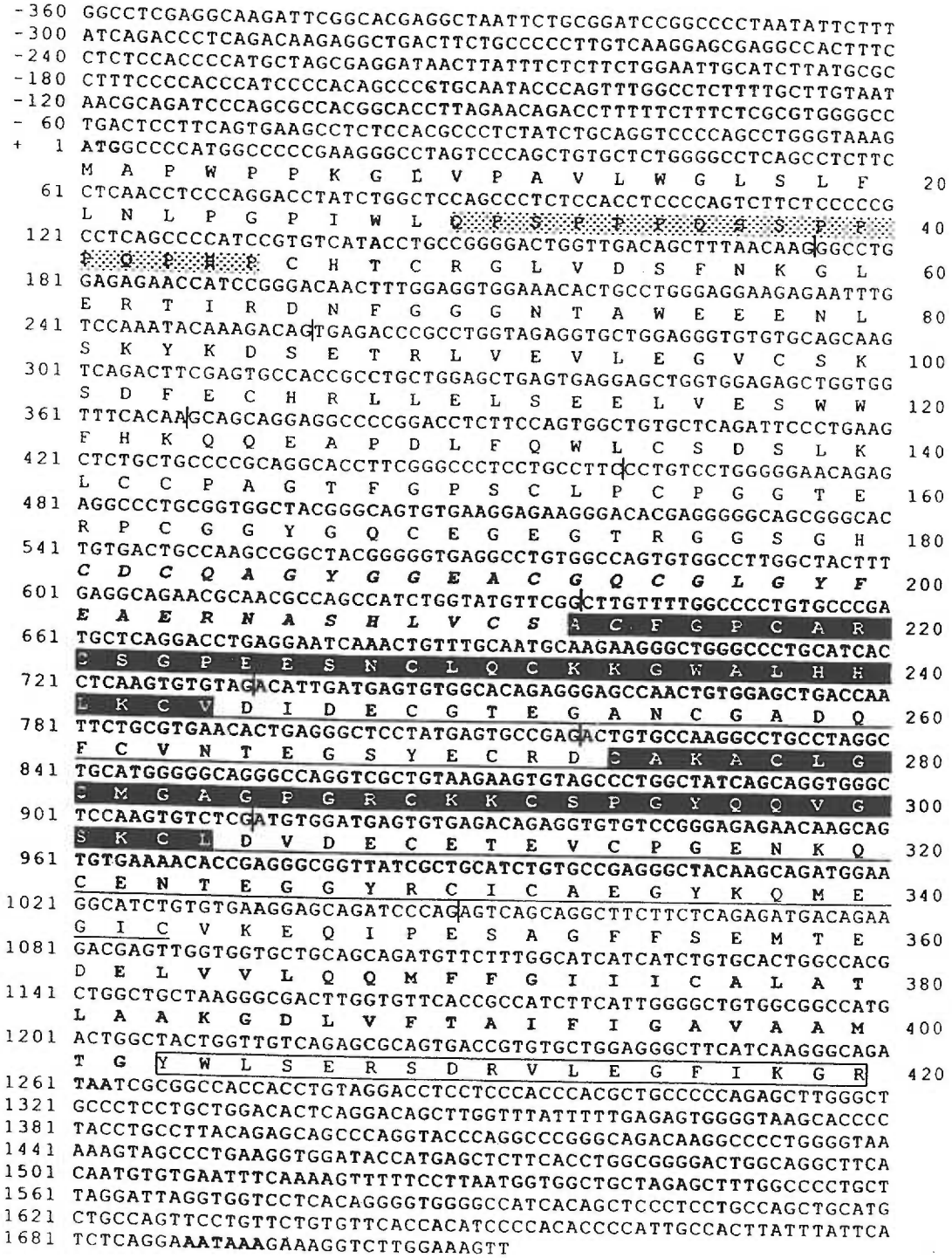
Chapter 3

Results

Characterization of cirrin cDNA

A screen of expressed sequence tag (EST) clones mapping to human chromosome 3p24.2-25 revealed a cDNA with sequence similarities to the fibrillins and other ECM proteins (Timmers, Whitney et al. 1996). This cDNA clone was selected for further characterization based on the hypothesis that ECM proteins play an important role in cardiovascular development. The clone was obtained from the Whitehead Institute (accession number WI-11041) and sequenced in its entirety. The sequence was confirmed from normal human fibroblast cDNA and normal genomic DNA. The cDNA clone is 2077 bases with the largest open reading frame being 1263 nucleotides (Figure 5 and Appendix 2, page 83). At the proposed translation-initiating methionine, a Kozak consensus sequence with a "strong context" for being the translational start site was found (Kozak 1984; Kozak 1996). There are five stop codons in frame upstream from the proposed translation-initiating methionine. Additional sequences upstream are indicative of an eukaryotic translational start domain (Ganoza and Louis 1994). The 5' and 3' untranslated regions (UTRs) are 365 and 449 bases, respectively. A termination codon begins at nucleotide 1261 and a polyadenylation signal is located at nucleotide 1689.

Figure 5. The complete cDNA sequence and amino acid translation for the cirrin gene. Nucleotide numbering is along the left border, numbering for the amino acid sequence is along the right border. For the DNA sequence, the initiation and termination codons, and the polyadenylation signal sequences are in bold typeface. Intron-exon boundaries are marked by vertical lines between the bordering nucleotides. For the amino acid sequence, the proline rich domain follows the cleavable secretion signal (a.a. 1-29) and is in the light shaded box. The laminin EGF-like domain is in bold italic typeface. The calcium binding 4-cys domain and the cbEGF-like domains are underlined. The furin-like cysteine rich repeats are in the dark shaded boxes. Note that there are two amino acid overlaps between the calcium binding domains and their adjacent furin-like cysteine rich repeats. The amino acids constituting a predicted transmembrane domain are in bold typeface. The predicted cytoplasmic domain is in the unshaded box.


-360 GGCCTCGAGGCAAGATTCGGCACGAGGCTAATTCTGCGGATCCGGCCCCTAATATTCTTT
-300 ATCAGACCCTCAGACAAGAGGCTGACTTCTGCCCCCTTGTCAAGGAGCGAGGCCACTTTC
-240 CTCTCCACCCCATGCTAGCGAGGATAACTTATTTCTCTTCTGGAATTGCATCTTATGCGC
-180 CTTTCCCCACCCATCCCCACAGCCCCTGCAATACCCAGTTTGGCCTCTTTTGCTTGTAAT
-120 AACGCAGATCCCAGCGCCACGGCACCTTAGAACAGACCTTTTTCTTTCTCGCGTGGGGCC
- 60 TGACTCCTTCAGTGAAGCCTCTCCACGCCCTCTATCTGCAGGTCCCCAGCCTGGGTAAAG
+ 1 ATGGCCCCATGGCCCCCGAAGGGCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTC
M A P W P P K G L V P A V L W G L S L F 20
61 CTCAACCTCCCAGGACCTATCTGGCTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCG
L N L P G P I W L Q P S P P P Q S S P P 40
121 CCTCAGCCCCATCCGTGTCATACCTGCCGGGGACTGGTTGACAGCTTTAACAAGGGCCTG
P Q P H P C H T C R G L V D S F N K G L 60
181 GAGAGAACCATCCGGGACAACTTTGGAGGTGGAAACACTGCCTGGGAGGAAGAGAATTTG
E R T I R D N F G G G N T A W E E E N L 80
241 TCCAAATACAAAGACAGTGAGACCCGCCTGGTAGAGGTGCTGGAGGGTGTGTGCAGCAAG
S K Y K D S E T R L V E V L E G V C S K 100
301 TCAGACTTCGAGTGCCACCGCCTGCTGGAGCTGAGTGAGGAGCTGGTGGAGAGCTGGTGG
S D F E C H R L L E L S E E L V E S W W 120
361 TTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCCAGTGGCTGTGCTCAGATTCCCTGAAG
F H K Q Q E A P D L F Q W L C S D S L K 140
421 CTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGCCTTCCCTGTCCTGGGGGAACAGAG
L C C P A G T F G P S C L P C P G G T E 160
481 AGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGGGACACGAGGGGGCAGCGGGCAC
R P C G G Y G Q C E G E G T R G G S G H 180
541 TGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCCAGTGTGGCCTTGGCTACTTT
C D C Q A G Y G G E A C G Q C G L G Y F 200
601 GAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCGGCTTGTTTTGGCCCCTGTGCCCGA
E A E R N A S H L V C S A C F G P C A R 220
661 TGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGGCTGGGCCCTGCATCAC
C S G P E E S N C L Q C K K G W A L H H 240
721 CTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACTGTGGAGCTGACCAA
L K C V D I D E C G T E G A N C G A D Q 260
781 TTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAGGCCTGCCTAGGC
F C V N T E G S Y E C R D C A K A C L G 280
841 TGCATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCAGCAGGTGGGC
C M G A G P G R C K K C S P G Y Q Q V G 300
901 TCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGAACAAGCAG
S K C L D V D E C E T E V C P G E N K Q 320
961 TGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAGATGGAA
C E N T E G G Y R C I C A E G Y K Q M E 340
1021 GGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGACAGAA
G I C V K E Q I P E S A G F F S E M T E 360
1081 GACGAGTTGGTGGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCACG
D E L V V L Q Q M F F G I I I I C A L A T 380
1141 CTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
L A A K G D L V F T A I F I G A V A A M 400
1201 ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGA
T G Y W L S E R S D R V L E G F I K G R 420
1261 TAATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCT
1321 GCCCTCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCC
1381 TACCTGCCTTACAGAGCAGCCCAGGTACCCAGGCCCGGGCAGACAAGGCCCCTGGGGTAA
1441 AAAGTAGCCCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCA
1501 CAATGTGTGAATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCTGCT
1561 TAGGATTAGGTGGTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATG
1621 CTGCCAGTTCCTGTTCTGTGTTCACCACATCCCCACACCCCATTGCCACTTATTTATTCA
1681 TCTCAGGAAATAAAGAAAGGTCTTGGAAAGTT

Genomic organization

The intron-exon structure was determined by DNA sequence analysis of PCR amplified fragments from BAC 172I17 and confirmed by sequence analysis of PCR amplified fragments of normal genomic DNA. There are ten identified coding exons, encompassing approximately 12 kb of genomic DNA (sequence can be found in Appendix 2, pages 84-85). The introns have been sequenced in their entirety, with the exception of introns-2, 4 and 6. In those cases the intron-exon boundaries and several hundred internal bases have been sequenced. All intron-exon boundaries have the appropriate splice site sequences and all intron sequences have polypyrimidine tracts and branchpoint consensus sequences characteristic of mammalian introns (Table 2). Figure 6 shows the genomic organization, with the sizes of the introns and exons indicated. The positions of the intron-exon boundaries are shown in Figure 5.

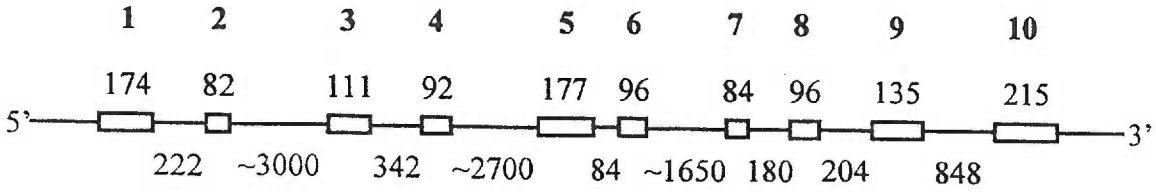


Figure 6. Diagram of the intron-exon boundary structure of the cirrin gene. The solid lines represent introns, the open boxes are exons. The exon numbers are in bold above the exons. The sizes, in bases, are below the corresponding introns and above the corresponding exons.

Table 2. Intron-exon splice sites.

Intron	Donor Site	Acceptor Site
1	CAAG gtgg	acttagctattactaattttctgtttccag GGCC
2	ACAG gtaa	tttcccaccagccctgccctgtccgatcag TGAG
3	ACAA gtga	cacctccctccaccctgcccctgccatcag GCAG
4	CTTC gtga	gacctcacctggtttggtgtcttcccacag CCTG
5	TCGG gtag	ccatcctcatgctgcccccattccacccag CTTG
6	GTAG gtaa	gaaattctcaccctgctcacctctctgcag ACAT
7	CGAG gtca	ctcaccctcatctttctctcctctctccag ACTG
8	CTCG gtga	gcaggactctgacccctccctcccctcaag ATGT
9	CCAG gtga	tgccaggctgcatctcttgctcctctgcag AGTC

Analysis of cirrin transcripts

Northern blot analysis of mRNAs from multiple human tissues was done to determine the transcript size and pattern of expression. Fetal and adult multiple tissue northern blots (Clontech) were hybridized with a cDNA probe incorporating the entire coding sequence (Figure 7). A transcript of 2.1 kb was observed in all tissues analyzed with the most prominent expression seen in fetal lung, liver, kidney and adult heart, brain and skeletal muscle. This transcript is consistent in size with the cDNA clone characterized here, confirming that it represents the full-length cDNA. Interestingly, the adult brain shows an alternative transcript of 2.5 kb that is not present in any other tissue examined including fetal brain. The larger transcript is detected when hybridized with a probe from the 5' half, but not the 3' half, of the gene indicating that the similarity of the transcripts lies in the 5' portion of cirrin. Whether this is an alternatively spliced transcript or a closely related gene product is unknown, but is currently under investigation. An RNA master blot (Clontech) and a northern containing mRNA from mouse, normal human skin fibroblasts, HT1080, WI138 and MG63 cell lines were also probed for cirrin. An appropriately sized transcript was present in all lanes (not shown).

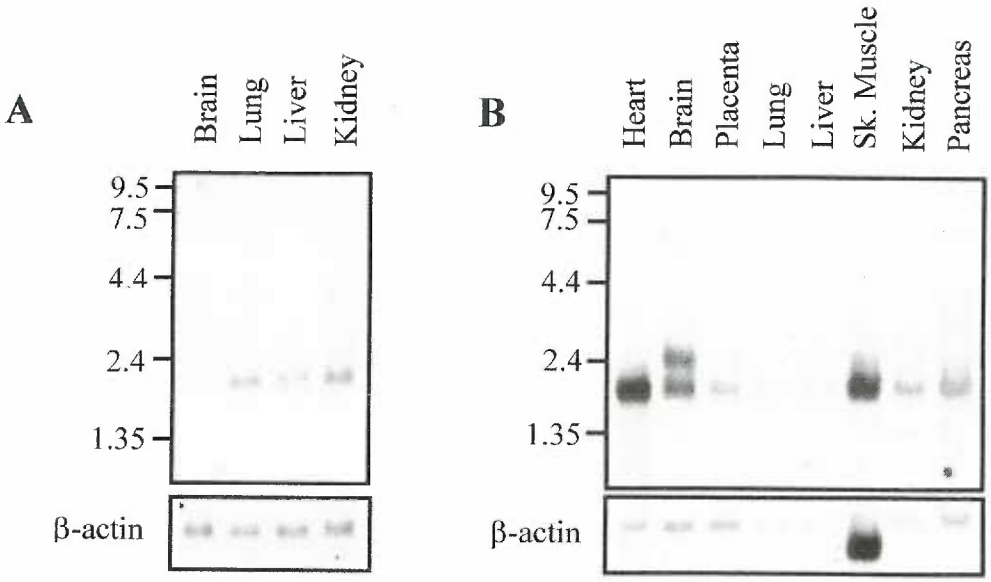


Figure 7. Northern blot analysis of poly A+ RNA from human fetal and adult tissues. (**A**) Fetal tissue northern blot hybridized with a cirrin specific cDNA probe encompassing the entire coding sequence. (**B**) Adult tissue northern blot, also hybridized with the cirrin probe. The tissue source is indicated above each lane. Note the 2.1 kb transcript in all lanes. Only adult brain expresses a second transcript of 2.5 kb. Below each blot is the same blot hybridized with a human β-actin cDNA probe used as a control to access the relative amounts of RNA present in each lane. Note that the β-actin for adult skeletal muscle appears to be more intense than the signal for other tissues, indicating that the signal for skeletal muscle cirrin is likely over-represented.

Whole mount *in situ* hybridization

To further determine the expression pattern of cirrin, whole mount *in situ* hybridization was performed on chick embryos of varying developmental stages. Ubiquitous staining was observed throughout embryos at HH stage 17 (Figure. 8A). By HH stage 22, high levels of expression can be seen in the developing heart, limb buds, mandible, branchial arches, brain, and around the somites and neural tube (Figure 8B). The most prominent staining in the heart was observed in the myocardium and endocardial cushions (Figure 8C and 8D), suggesting a possible role in cardiac development. A negative control is shown in Figure 8E. Hamburger and Hamilton (HH) staging is reviewed in Table 1.

Protein sequence analysis

Cirrin encodes a putative 420 amino acid protein with regions of similarity to elements present in extracellular matrix proteins. The deduced amino acid sequence is shown in Figure 5 (also Appendix 2, page 86). A diagrammatic representation of cirrin is found in Figure 9A. Recognizable structural motifs include a proline rich domain, an EGF-like domain similar to those found in the laminin family and a calcium binding epidermal growth factor-like domain (cb-EGF) with greatest similarity to the tandemly repeated cbEGF domains in the fibrillins. In addition, a four-cysteine domain (cb-4cys) with the calcium binding consensus sequence is present. Surprisingly, cirrin also contains two furin-like cysteine rich repeats that have not been found in other non-furin proteins. Regions of similarity can be observed in Figure 9B. There is also a unique tryptophan and glutamic acid rich (WE) region. Exons 2-4 code for a protein domain that is rich in tryptophan and glutamic acid, 4.2% and 13.8% respectively.

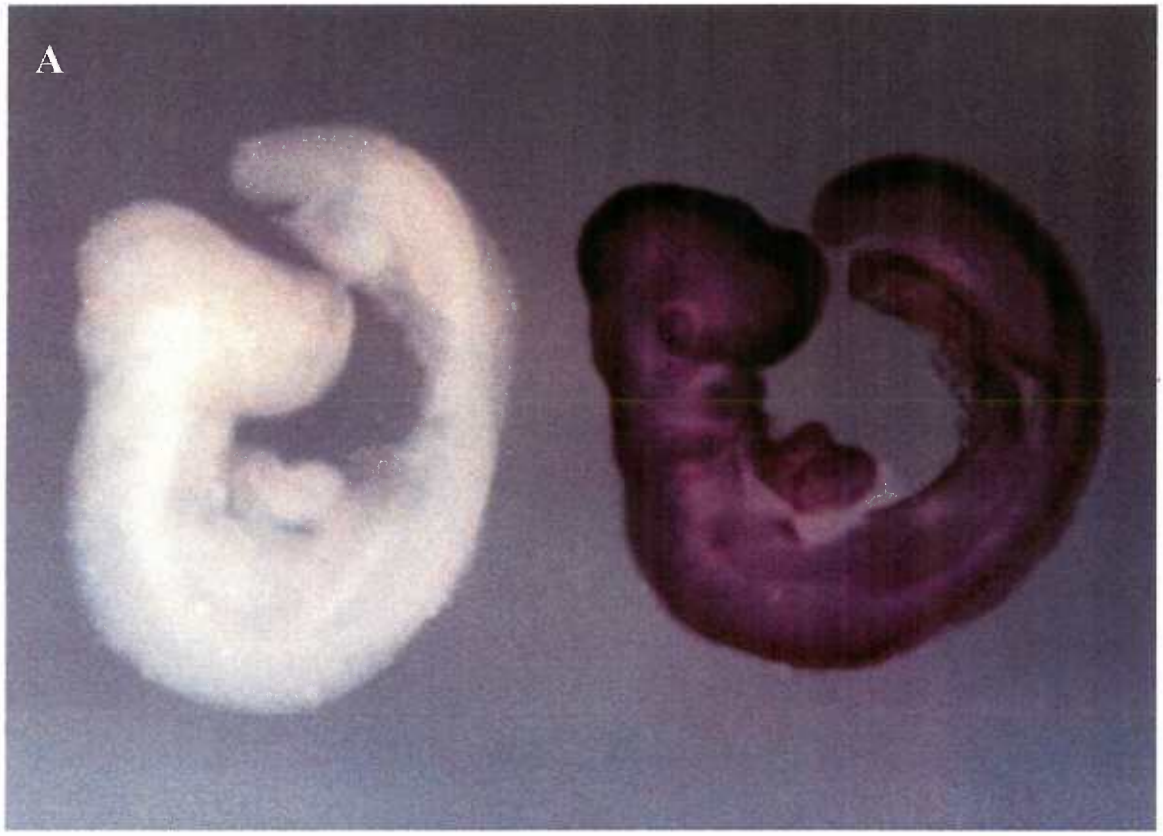
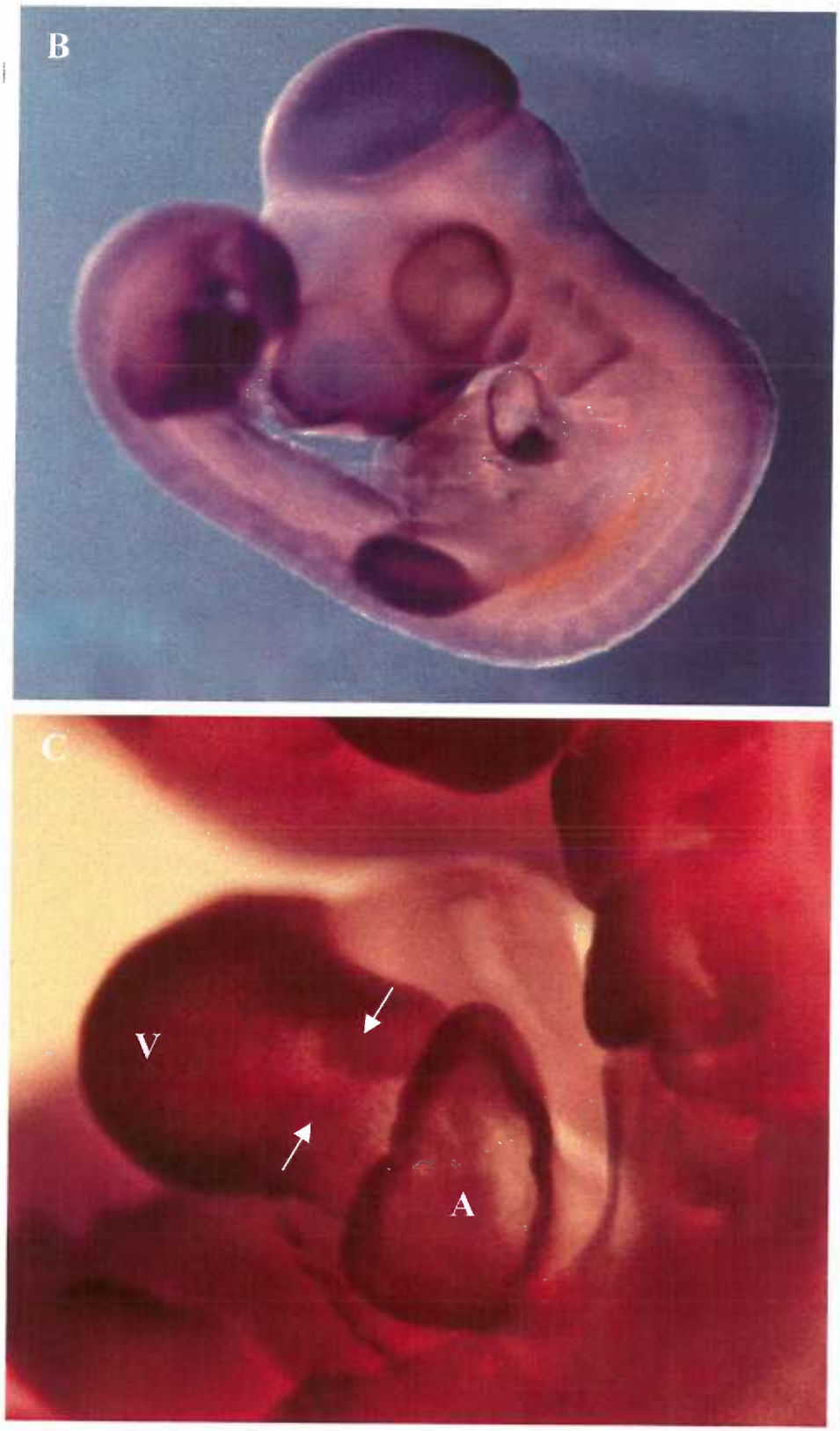
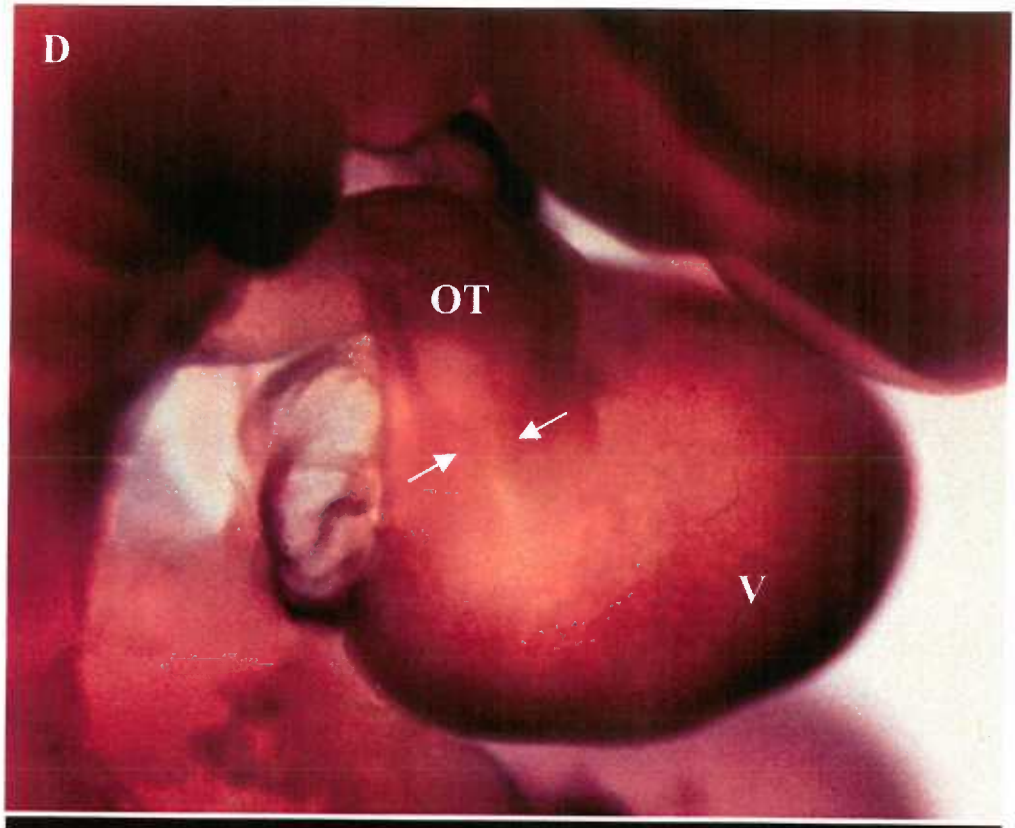
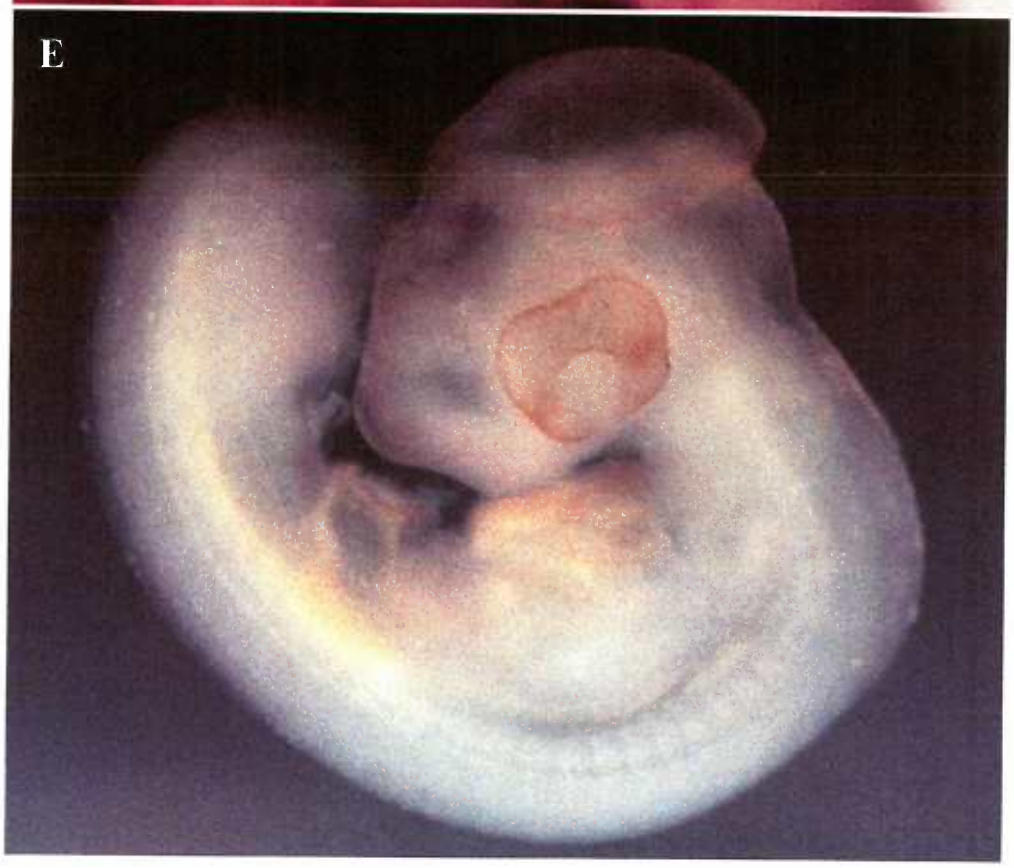


Figure 8. Whole-mount *in situ* hybridizations of chick embryos. (**A**) Whole chick embryos after *in situ* hybridization with a sense (left) and antisense (right) cirrin-specific RNA probe. Note ubiquitous staining. Right view, HH stage 17, magnification 11x. (**B**) Embryo after *in situ* hybridization with cirrin-specific antisense probe. Note staining in developing brain, limb buds, cardiac atrial muscle and cushion tissue, branchial arches and around somites. Left view, HH stage 22, magnification 12.5x. (**C**) Close view of chick embryo heart from the left side following *in situ* hybridization. Note expression in atrial and ventricular myocardium as well as cushion tissue (arrows). A, atrium; V, ventricle. HH stage 22, magnification 32x. (**D**) Close view of chick embryo heart from the right side following *in situ* hybridization with antisense probe. Note expression in outflow tract and cushions (arrows). OT, outflow tract; V, ventricle. HH stage 22, magnification 27x. (**E**) Chick embryo after *in situ* hybridization with the sense strand from the cirrin-specific probe as a negative control. Note the absence of any staining. Left view, HH stage 22, magnification 12x.


B
C
V
A


D
OT
V


E

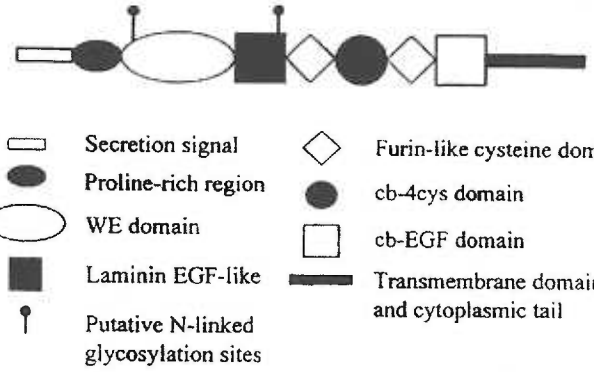
If looking only at exons 2 and 3, the putative protein domain contains 4.8% tryptophan and 19.0% glutamic acid. Overall, cirrin is 2.1% tryptophan and 9.3% glutamic acid rich.

In addition to recognizable protein domains, other sequence-based elements are indicated. There are two consensus sites for N-linked glycosylation, one in the WE domain and one in the laminin EGF-like domain. The calcium binding domains are recognized by the β-hydroxylation consensus sequence, [(D/N)X(D/N)(Q/E)X_n(D*/N*)X_m(Y/F); * indicates potential hydroxylation], which is associated with calcium binding in other proteins including the fibrillins (Glanville, Qian et al. 1994). Computer analyses also predict a secretion signal sequence at the amino-terminus with a cleavage site between amino acid residues 29-30, and a type I transmembrane domain between amino acid residues 362-402. This model suggests that cirrin has an 18 amino acid cytoplasmic domain at the carboxyl-terminus of the molecule, with the bulk of the molecule (amino acids 1-361) residing in the extracellular space.

Protein characterization

Western blot analysis of protein produced in a baculovirus expression system using affinity purified polyclonal antibody 1851 recognizes a single band with an apparent molecular weight of approximately 55 kDa that is present only in cells infected with the cirrin-baculovirus construct (Figure 10A). This interaction can be inhibited by pre-incubating the antibody with excess free peptide, demonstrating the specificity of the antibody for cirrin (Figure 10b). As a control, a baculovirus construct with no insert was used. A single 65 kDa band was also detected in an extract of fetal bovine cartilage (Figure 10A). This interaction was also abolished in the competitive inhibition assay

A



B

Proline-rich region:

```
cirrin          30 QPSPPPQSSPPPQPHP 45
                   QP+PPP   P PQP P
Fbn2               QPQPPPPKPPRPQPPP
```

Laminin EGF-like domain:

```
cirrin         181 CDCQAGYGGEACGQCGLGYFEAERNASHLVCSAC 214
                   C+C  GY G +C  C  GY+  +       +C  C
Laminin-α1         CECPQGYTGTSCEACLPGYYRVDGILFGGICQPC
```

Calcium binding EGF-like domains:

```
cirrin         245 DIDECGTEGANCGADQFCVNTEGSYEC 271
                   DI EC+    NCG   +C NT GS+ C
Fbn1               DINECIGAH NCGKHAVCTNTAGSFKC

cirrin         305 DVDECETEV  CPGENKQCENTEGGYRCICAEGYK Q MEGIC 343
                   DVDEC + +  C  ++  C NT G+YRC C  GY        C
Fbn1               DIDECQNGPV CQRNAE CINTAGSYRCDCKPGYRFTSTGQ C
Fbn1               DVDECAENINLCENGQ  CLNVPGAYRCECEMGFTPASDSRSC
Fbn2               DVDECAENINLCENGQ  CLNVPGAYRCECEMGFTPASDSRSC
LTBP2              DIDECANDTM CGSHGF CDNTDGSFRCLCDQGFEISPSGWDC
```

Furin-like cysteine rich repeats:

```
cirrin         214 CFGPCARCSGPEESNCLQCKKGWALHHLKCV 244
                   C++ C  CSGP  ++C+QC  GW L  + CV
Furin2             CYLSCHTCSGPRRNQCVQCPAGWQLAAGECV
                   C+  C  C+G     C  C  G+Q ++  C+
cirrin         274 CAKACLGCMGAGPGRCKKCSPGYQQVGSKCL 304
                   C + CL C GA+   C  CSPG Q     C
Furin1             CDRSCLECYGALASQCSTCSPGSQ     C
```

Figure 9. (A) Diagrammatic representation of cirrin showing the predicted domain structure. The symbols representing the various domains are identified under the diagram. (B) Alignment of cirrin domains with similarity to known protein domains.

(Figure 10B), indicating that the 65 kDa band is fetal bovine cirrin. Western blot analysis of protein precipitated out of conditioned medium from cultured human fibroblasts detects a major band with an apparent molecular weight of about 65 kDa, with a faint lower band present (Figure 10C). The difference between the apparent molecular weights of cirrin, those produced in a recombinant baculovirus system, extracted from bovine cartilage and isolated from cultured normal fibroblast media, and the predicted molecular weight of 45 kDa, may be due to the presence of EGF-like domains, which have been shown to alter the electrophoretic mobility of proteins on SDS-PAGE (Persson, Selander et al. 1989; Rand, Lindblom et al. 1997). Cirrin is a very cysteine rich protein (9%), with most cysteines putatively involved in internal disulfide bonds. Running cirrin under reduced conditions unfolds the EGF-like domains and other cysteine rich domains possibly accounting for difference in predicted versus apparent molecular weights. Although the insect cells used in the baculovirus expression system are capable of producing many posttranslational modifications seen in mammalian proteins, native cirrin may have some modifications that are unable to be produced by these cells, thus accounting for differences between recombinant and native cirrin. Contrary to the sequence-based prediction that cirrin is a membrane bound protein, these data indicate that cirrin is secreted into the extracellular space. However, it is possible that there is also a membrane bound form of cirrin.

Conservation across species

Southern blot analysis of genomic DNA from multiple diverse species indicates that the cirrin gene is highly conserved (Figure 11). Hybridization of a cirrin-specific cDNA probe under stringent conditions detected cirrin-related sequences in all represented

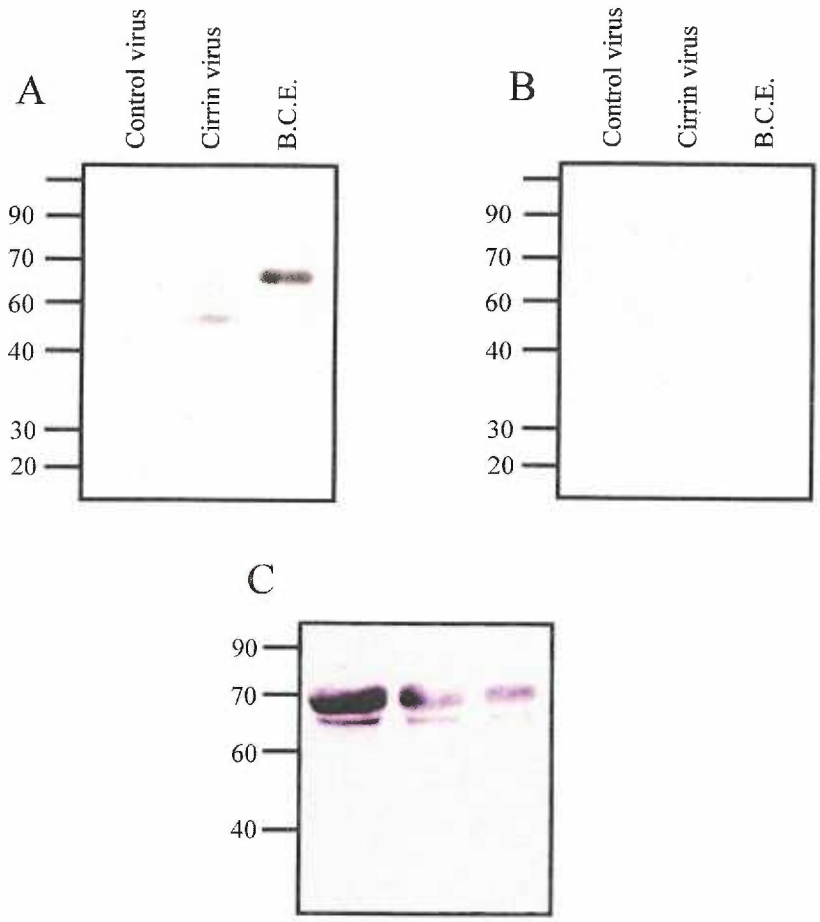


Figure 10. (**A**) Western blot of insect cell culture medium infected with non-recombinant baculovirus as a negative control (control virus), recombinant baculovirus expressing the cirrin gene (cirrin virus), and a low salt extract of fetal bovine cartilage (BCE). Note the single band of 55 kDa that is present in the medium from insect cells expressing cirrin, but absent in medium from insect cells infected with baculovirus that does not express an exogenous gene. There is a 65 kDa band detected in the bovine cartilage extract. (**B**) A duplicate western blot to that shown in panel A. The antibody was pre-incubated with the peptide antigen prior to incubation with the blot, blocking detection of both the 55 kDa and 65 kDa proteins. (**C**) Western blot of protein precipitated from human fibroblast cell culture medium. There are 3 lanes containing differing amounts of a single protein sample. In each lane there is a 65 kDa band similar to that seen in the bovine cartilage extract. A slightly smaller secondary band is also present possibly due to variation in glycosylation. Molecular weights in kDa.

species except yeast. BLAST searches of gene and protein databases revealed no significant matches with other human genes outside of the regional similarities with individual protein domains (Figure 8B). However, the amino acid sequence of cirrin does have high similarity (47% identical, 60% similar) overall to the predicted amino acid sequence for a gene from *Cricetulus griseus* (Chinese hamster). This gene encodes an uncharacterized putative extracellular protein named HT protein (Genbank accession number U48852). The coding region of cirrin overall is 44% identical (58% similar) with F09E8 from *C. elegans* (Genbank accession number Z73896) which also encodes a protein of unknown function. When looking specifically at the WE domain (cirrin amino acids 46-140) alignment, the identity increases to 56% for HT (73% similar). The WE domain and F09E8 are 53% identical (70% similar) over this same region. Alignment of the amino acid sequences for cirrin, HT and F09E8 is shown in Figure 12.

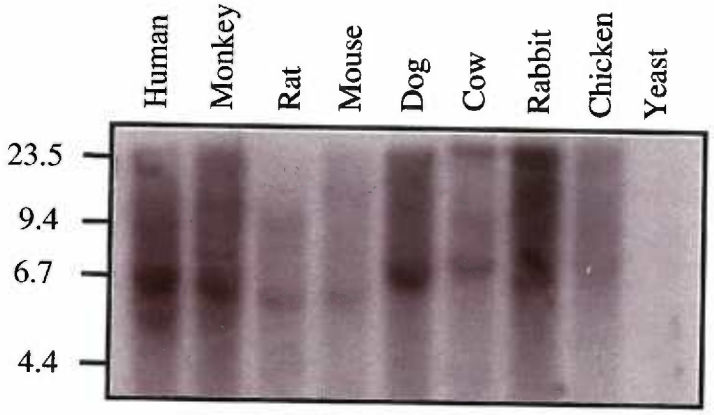


Figure 11. Southern blot analysis of genomic DNA from different species. The species represented are indicated above each lane. Sizes in kilobases (kb).

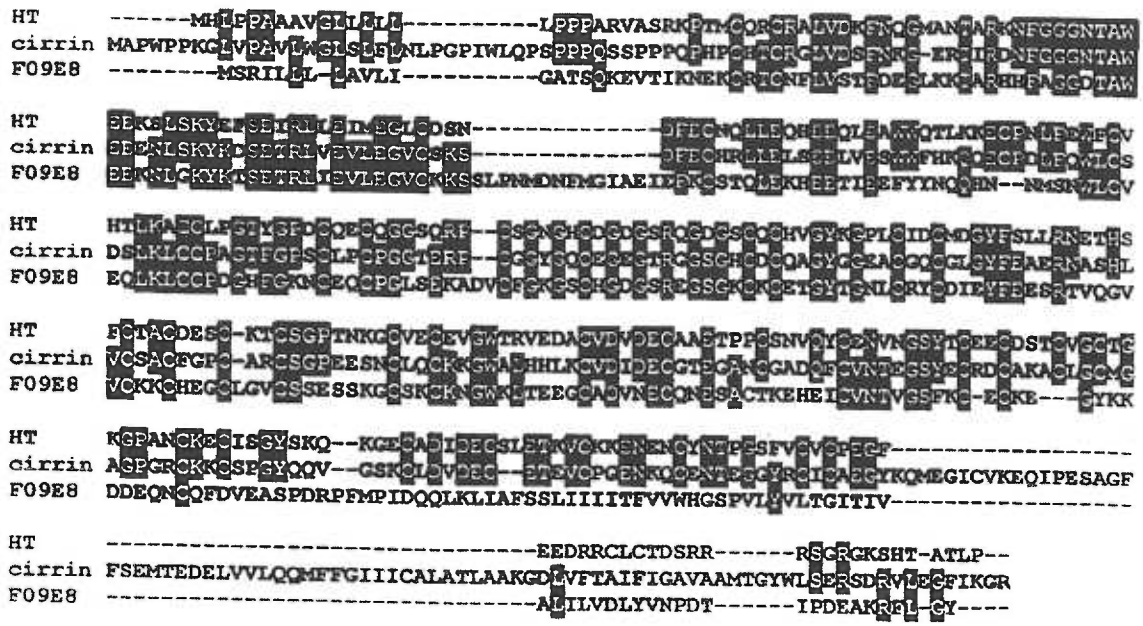


Figure 12. Alignment of cirrin with HT protein (Chinese hamster) and F09E8 (*C. elegans*). Overall, cirrin is 47% identical to the HT protein and 44% identical to F09E8.

Fluorescent *in situ* Hybridization (FISH)

Three well-characterized cell lines with terminal deletions of the short arm of chromosome 3 distal to band p25 (3p-) were obtained from the NIGMS Human Mutant Cell Repository (Coriell Institute). Two of the cell lines (GM10922 and GM07873) were from individuals with congenital heart defects. Cell line GM10985 was derived from a 3p- patient with no heart malformation. All cell lines originated from individuals who manifested characteristic phenotypic features of 3p- syndrome. FISH with BAC172I17 and a chromosome 3 alpha satellite probe was performed on metaphase chromosome spreads from all cell lines and a normal control. Figure 13 shows representative FISH results for GM10922 (deleted for cirrin) and a normal cell line. All control metaphase

cells demonstrated signals for the cirrin probe and the alpha satellite probe on both chromosome 3 homologues. In each chromosome spread examined from the lymphoblast line GM10922 and fibroblast line GM07873, only one probe signal for cirrin was present on the normal chromosome 3, while both homologues had the alpha satellite signal. In the lymphoblast cell line GM10985, four cells had a cirrin signal on both the normal chromosomes 3 and on the deleted chromosome 3; ten cells had only one signal on the normal chromosome. All 14 cells had two alpha satellite signals. Results are summarized in Table 3. The latter inconsistent hybridization pattern has been observed previously (personal communication, Dr. Susan Olson) when the probe target sequence is present adjacent to the breakpoint of a deleted chromosome, possibly reflecting a vulnerability to probe drop off. To try to eliminate this phenomenon, FISH analysis was performed on all 3p- cell lines and a normal control using an 11 kb cirrin-specific probe. Similar results were obtained for GM10922 and GM07873 in which only one cirrin signal was observed on the normal chromosome. With GM10985, two cells contained two signals while nine cells contained only one signal on the normal chromosome. All had two alpha satellite signals indicating two copies of chromosome 3 were present per metaphase spread (summarized in Table 3). To determine if there was mosaicism for the chromosome deletion, 50 cells from GM10985 were examined by G-banding and all contained one normal and one deleted chromosome 3.

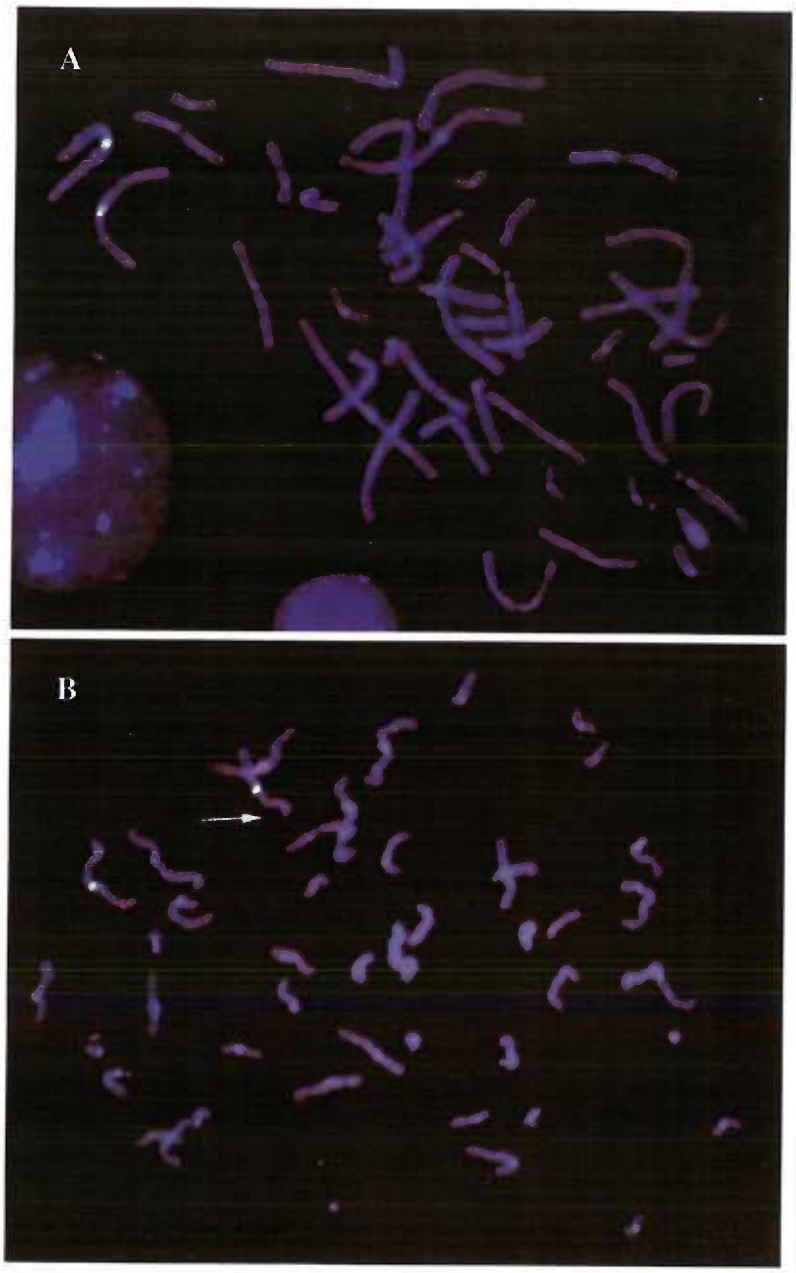


Figure 13. Fluorescent in situ hybridization with chromosome 3 alpha-satellite probe (yellow) as identifier and the cirrin probe (red) with DAPI counterstain. (A) Normal metaphase spread showing hybridization signals on both chromosome 3 homologues for 3 alpha-satellite and cirrin probes. (B) Metaphase spread from cell line GM10922 [46,XY,del(3)(p25)] showing hybridization signals on both copies of chromosome 3 for the alpha-satellite identifier probe; however, the cirrin probe signal is absent on the deleted chromosome 3 (arrow) and present on the normal chromosome 3.

Table 3. Summary of 3p- FISH analysis.

Cell line	BAC 172I17		11 kb cirrin probe		Total	
	1 signal	2 signals	1 signal	2 signals	1 signal	2 signals
GM07873	10	0	9	0	19	0
GM10922	14	0	20	0	34	0
GM10985	10	4	9	2	19	6
Normal	0	all	0	all	0	all

Southern blot analysis of GM10922 and GM10985

To determine if GM10985 had a deletion within the cirrin gene, possibly accounting for the inconsistent hybridization observed during FISH analysis, Southern blot analysis was employed. Genomic DNA from GM10985, GM10922 and normal controls was digested with PstI or EcoRI, transferred to a nylon membrane and probed with the complete coding region of cirrin. DNA from 20 normal unrelated copies of chromosome 3 gave six bands with PstI restriction, while EcoRI restriction gave two bands (8 normal chromosomes). No junctional fragments were observed in GM10985, suggesting that the cirrin gene is either intact on both copies of chromosome 3 or completely missing on the deleted chromosome 3. Restriction of GM10922 with PstI (deleted by FISH analysis) resulted in the presence of a unique ~1 kb junctional fragment (Figure 14). A unique junctional fragment was also observed with EcoRI restriction (not shown), suggesting GM10922's chromosomal breakpoint lies within the cirrin gene. Additional enzymes will be looked at.

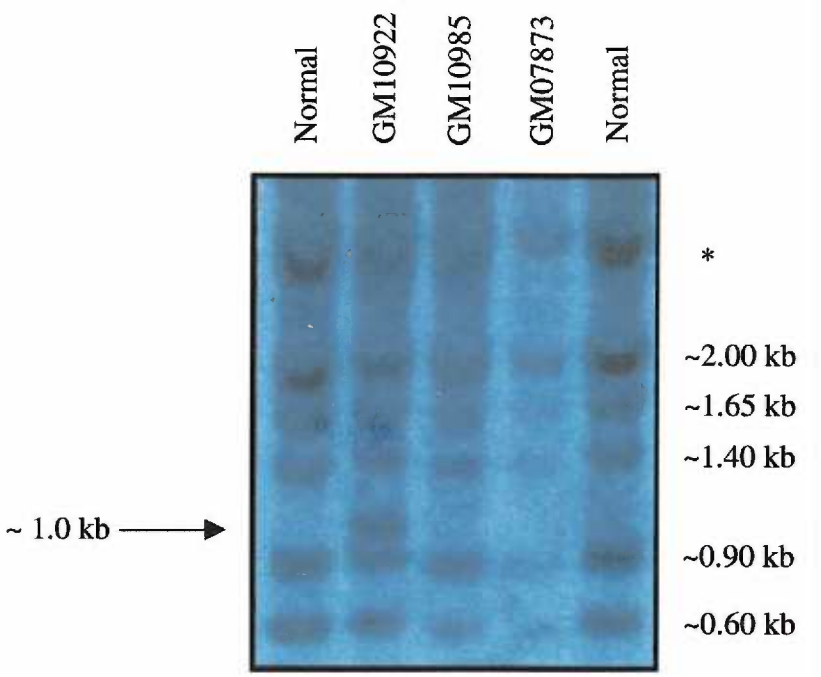


Figure 14. Southern blot of PstI digest on 3p- cell lines. Approximate sizes of normal bands are indicated on the right. * indicates a band larger than the marker used and therefore the size is undeterminable. Unique ~1.0 kb junctional fragment in GM10922 is indicated by the arrow.

Chapter 4

Discussion

The formation of a normal functioning four-chambered heart is a complex process of cell signaling, migration and differentiation, cell-cell interactions, and cell-matrix interactions. Alterations of any of the numerous steps of cardiogenesis are likely to result in cardiac defects, the majority of which are caused by the improper formation of valves and membranous septa (Potts, Dagle et al. 1991). Mechanistically, some defects in valves and septa are attributed to malformations in the extracellular matrix (Clark 1995). The ECM plays a prominent role in the formation of the endocardial cushions required for proper septation, valve leaflet formation, and structural alignment of the heart. Consequently, genes expressed in these areas during cardiac morphogenesis are prime candidates for cardiac septal defects as well as other heart malformations. This work describes the partial characterization of a new gene and its protein product named cirrin. The evidence presents cirrin as a very good candidate gene for heart malformations sometimes associated with 3p- syndrome and we propose that mutated cirrin is a potential candidate for non-syndromic cardiac defects.

Structure of cirrin gene

The cirrin gene has an open reading frame of 1263 bases within a 2.1 kb transcript. An alternative transcript of 2.5 kb can be found in the adult brain. It is not known whether this is an alternatively spliced gene product or closely related gene, but it is currently being investigated. At the genomic level, cirrin is comprised of 10 exons spanning ≈12 Kb. The appropriate consensus sequence for all splice donor and acceptor sites were found.

Structure of cirrin protein

At this point, the role of cirrin is unknown and predicting a function can only be speculated. However, the predicted protein structure provides some information with regard to possible function. Cirrin has putative structural features found in many extracellular matrix proteins, such as a calcium binding EGF-like domain and a proline rich region with high similarity to those found in the fibrillins. A laminin EGF-like domain is also present. EGF-like domains are common structural features of many extracellular proteins having a variety of functions. They are found in proteins involved in blood clotting, such as factor VII, factor IX, factor X, protein C, and factor XII (Furie and Furie 1988). EGF-like domains are also found in adhesion molecules such as P-selectin. This multidomain adhesion protein, found on the surface of activated platelets and endothelial cells that functions in the recruitment of leukocytes to sites of inflammation, requires both a lectin and an EGF-like domain for optimal ligand binding (Freedman, Sanford et al. 1996). Molecules involved in signaling, such as the Notch receptors and their corresponding ligands, also contain EGF-like domains (Artavanis-Tskonas, Matsuno et al. 1995; Kimble, Henderson et al. 1998). The EGF-like domains in the Notch receptors are involved in receptor dimerization and are required for receptor-ligand binding (Hartley, Xu et al. 1987; Kimble, Henderson et al. 1998). In contrast, the EGF-like domains in the Notch ligands are not essential for signaling and their function is as yet unknown (Kimble, Henderson et al. 1998). A number of ECM proteins also contain EGF-like domains, such as laminin (Mayer, Nischt et al. 1993), tenascin/cytoactin (Gulcher, Nies et al. 1989), thrombospondin (Engel 1989), fibulin (Argraves, Tran et al. 1990), latent transforming growth factor-β binding protein-1

(LTBP-1) (Kanzaki, Olafsson et al. 1990) and fibrillin (Maslen, Corson et al. 1991) to name a few. A single EGF-like domain in the basement membrane protein laminin is known to bind nidogen, another molecule containing EGF-like domains (Mayer, Nischt et al. 1993). The glycoprotein fibulin-1 is capable of interacting with itself and fibronectin through its fifth and sixth EGF-like domains (Tran, Dusen et al. 1997). Some EGF-like domains contain putative calcium binding sites that are recognized by the β-hydroxylation consensus sequence, [(D/N)X(D/N)(Q/E)X_n(D*/N*)X_m(Y/F); * indicates potential hydroxylation] (Glanville, Qian et al. 1994). Binding of Ca^{2+} to cbEGF domains is important in the structural stabilization and proper function of these regions (Engel 1989; Engel 1990; Maurer, Mayer et al. 1992; Handford, Downing et al. 1995; Maurer and Hohenester 1997; Reinhardt, Mechling et al. 1997; Reinhardt, Ono et al. 1997; Yuan, Downing et al. 1997; Cardy and Handford 1998). In addition to the cbEGF-like domain, cirrin has a four-cysteine domain (cb-4cys) which also contains the β-hydroxylation consensus sequence.

Proline rich domains are found in a wide variety of structurally and functionally unrelated proteins where they commonly act as sites of intermolecular interactions. These domains exist as repetitive short proline-rich regions, i.e. $(XP)_n$ or $(XPY)_n$, or in multiple tandem repeats with minor variations between repeated sequences (Williamson 1994). The light chain myosin kinase has the repetitive sequence $(AP)_6$ at its amino-terminus where it contributes to binding of actin (Frank and Weeds 1974). NMR studies of the $(AP)_6$ region of the light chain myosin kinase indicate it can be viewed as an elongated arm extending away from the rest of the molecule (Bhandari, Levine et al. 1986). Another protein with a proposed rigid arm extending from the bulk of the

molecule is the RNA polymerase II. The carboxy-terminal domain (CTD) is composed of a heptapeptide repeat (YSPTSPS) that varies from 26-52 copies in yeast to mammals (Koleske, Buratowski et al. 1992; Usheva, Maldonado et al. 1992). The CTD is proposed to bind to the transcription factor TFIID as one of the initial steps in the formation of a transcription-competent complex (Koleske, Buratowski et al. 1992; Usheva, Maldonado et al. 1992). Other proteins have non-repetitive proline-rich regions that are involved in protein-protein interactions. Probably the most well-known, are those proteins that interact with the Src homology 3 (SH3) domain present in a very large group of proteins, including signaling proteins (Crk, Grb2) and cytoskeletal elements (spectrin, myosin) (Ren, Mayer et al. 1993; Rozakis-Adcock, Fernley et al. 1993). Ren and coworkers determined that the SH3 binding sites on 3BP1 and 3BP2, which bind the Abl proto-oncogene product, contained the consensus sequence XPXXPPPψXP (ψ represents hydrophobic amino acid residue) (Ren, Mayer et al. 1993). Similarly, it was found that the SH3 domain of Grb2 bound to the proline-rich carboxy-terminal tail of mSos1, a protein required for Ras signaling (Rozakis-Adcock, Fernley et al. 1993). In looking at the proline-rich region of cirrin, the repetitive sequence XP can be found in the form of $(XP)_3X_2(XP)_4$.

In addition to domains found in ECM proteins, there are two furin-like cysteine rich repeats in cirrin. Furin is a member of a calcium-dependent serine protease family homologous to the yeast propeptidase Kex2 and bacterial subtilisin. This particular type of repeat has not been found in any other non-furin protein. The function of these repeats in the furins is unknown, although they apparently are not essential for enzymatic activity (Hatsuzawa, Murakami et al. 1992). However, conservation of these repeats in various

furin family members suggests that there is a functional role that is yet to be defined (Roebroek, Creemers et al. 1992; Nakagawa, Murakami et al. 1993). As there is no other sequence similarity between cirrin and the furin family of subtilisin-like endoproteases, and in particular no evidence of a furin-like catalytic domain, it is clear that cirrin is not a member of that protein family.

The unique tryptophan and glutamic acid rich (WE) region has no known human homologous sequences. However, it is a highly conserved domain with 56% amino acid identity to a similar region in the Chinese hamster HT protein, and a surprising 53% identity to part of the coding region of the *C. elegans* gene designated F09E8. Both the HT protein and the product of F09E8 are uncharacterized, although the HT protein is predicted to be extracellular. The unusual composition and the highly conserved nature of the WE domain suggests that it plays a role in cirrin function. Glutamic acid rich regions are predicted to assume a coiled coil formation and are often involved in binding other molecules (Scartezzini, Egeo et al. 1997; Lee, Kambe et al. 1998). They have been detected in proteins of gene regulatory multiprotein complexes, such as transcriptional activators, α and β-tubulin, and G protein β subunits (Lupas, Van Dyke et al. 1991; Degan, Agterbos et al. 1999).

Data presented here demonstrate that cirrin is secreted into the medium of cultured fibroblasts, and can be detected in extracts of fetal bovine cartilage, indicating that it is an extracellular protein. However, prediction of a carboxy-terminal transmembrane domain by analysis of the cDNA sequence suggests that there may also be a cell membrane bound form, with a short cytoplasmic tail. There is a growing list of proteins that are grouped by their ability to exist in both an insoluble membrane-bound

form and a soluble secreted form. Generation of these two forms can occur by different biosynthetic pathways (alternative splicing or closely related but distinct genes) or by release of an extracellular domain from a membrane bound form. The second method of producing a soluble secreted molecule often involves type I transmembrane proteins that have a cleaved signal peptide at the N-terminus followed by an extracellular domain of variable size that is usually glycosylated and contains structural motifs, a single membrane-spanning domain, and a typically smaller intracellular domain (Slentz-Kesler, Hale et al. 1998). This second method fits the description of cirrin perfectly. Examples of proteolytically released proteins are numerous, and include cytokines (TGF-α, TNF-α, CSF-1), cytokine receptors (receptors for TNF,CSF-1, IL-2, IL-4), leukocyte antigens (CD8, class I MHC, CD14), ectoenzymes (ACE, sialyltransferase), and cell adhesion molecules (Mel-14, ELAM-1, NCAM) (Ehlers and Riordan 1991; Slentz-Kesler, Hale et al. 1998).

The cirrin protein appears to be a series of potential protein binding domains linked in series. What these domains interact with is food for future work.

Possible functions

The term "adheron" was first coined by Schubert and LaCorbiere to describe multicomponent protein complexes, composed of fibronectin, collagen and several glycoproteins, which were isolated from myoblasts and appeared to have a biological function in cell-cell or cell-substratum adhesion in culture (Schubert and LaCorbiere 1980). The term was adapted for use to describe the particulate matrix complexes involved in signaling endocardial cells to undergo a transition to mesenchyme in the

AV/OT regions of the developing heart. A polyclonal antiserum (ES1) was prepared against EDTA soluble extracellular proteins extracted from embryonic chick hearts and found to recognize components of the adheron (Krug, Runyan et al. 1987). Interestingly, ES1 antigens are not restricted to the heart and can be found in other sites involved in inductive interactions, including limb mesoderm and apical ectodermal ridge, and in the ECM of trunk neural crest formation, neural tube and notochord (Isokawa, Krug et al. 1991; Mjaatvedt, Krug et al. 1991). As an extracellular protein, it is possible that cirrin is a component of the adheron. In addition to being expressed in the endocardial cushions, cirrin is highly expressed in the developing limb bud, branchial arches, brain, neural tube and around the somites. Four major ES proteins of 27, 44, 63, and 70 kD were affinity purified by Isokawa and coworkers using a polyclonal antiserum (ES3) raised against EDTA extractable proteins from embryonic chick hearts (Isokawa, Rezaee et al. 1994). In the same study, it was determined that the 70 kD protein was transferrin, but the identity of the remaining proteins could not be determined due to heterogeneous sequence information. We have shown, under reducing conditions, cirrin migrates around 65 kD in SDS-PAGE. Of the ECM proteins known to be components of adherons, antibodies to hLAMP-1 and ES/130 have been shown to block the epithelial to mesenchyme transition (Rezaee, Isokawa et al. 1993; Sinning and Hewitt 1996) suggesting roles as signaling proteins. It is thought that fibronectin, via its multiple binding domains, can bind to other members of the adheron and signal transformation by delivering the biologically active myocardial component to the endocardium (Mjaatvedt, Lepera et al. 1987). Fibronectin is incapable of directing the transformation process itself (Mjaatvedt, Lepera et al. 1987). It is possible that cirrin, with all of its potential protein binding domains, may act like

fibronectin and help hold the multimeric protein complex together or help deliver the signal across the cardiac jelly.

In addition to the particulate matrix of the adheron, a fibrillar matrix exists in the endocardial cushions as well as other sites of inductive cell interactions. With regards to endocardial cushion tissues, the non-particulate matrix is known to contain laminin, heparin sulfated proteoglycans, hyaluronan, collagens (I, IV, VI), fibronectin, vitronectin, tenascin, fibulin-1, fibrillin-1, fibrillin-2, LTBP-1 and emilin at varying times during valve and septa formation (Funderburg and Markwald 1986; Kitten, Markwald et al. 1987; Hurle, Kitten et al. 1994; Bouchey, Argraves et al. 1996; Kitten, Kolker et al. 1996; Spicer, Augustine et al. 1996; Nakajima, Miyazono et al. 1997). Our whole mount *in situ* results indicate that cirrin is ubiquitously expressed in early embryos, with more specific expression seen in later staged embryos. However, cirrin does maintain a ubiquitous pattern of expression in the later staged (HH stage 22) embryos, albeit not as strong as younger embryos (HH stage 17 and under). Given this, and the predicted protein structure of cirrin, one could hypothesize that cirrin is a component of the ECM. Cirrin can be extracted from cartilage with a low salt wash, suggesting that it would probably be a matrix associated protein rather than a structural component of the ECM.

The existence of a putative transmembrane domain provides another alternative. It is possible that a membrane-bound form of cirrin could act as a cell adhesion molecule with an extracellular "binding" domain at the amino terminus anchored by the carboxy terminal transmembrane domain. Cleavage near the membrane would allow for a released soluble form.

FISH results

Cell lines GM07873 and GM10922, both with CHD, were found by FISH analysis to be hemizygous for the cirrin gene. Unexpectedly, GM10985 (no CHD) gave inconsistent hybridization in which some cells contained one signal (normal chromosome 3) and others contained two signals (normal and deleted chromosomes). When combining FISH data from both the 11 Kb cirrin specific probe and the 100 Kb BAC 172I17 probe (contains cirrin), three times as many cells contained one signal as compared to two signals. The presence of two chromosome 3 alpha satellite signals in all cells examined, demonstrates that cirrin was not absent due to a missing chromosome 3. To rule out the possibility of mosaicism within the cells, 50 cells from GM10985 were examined by G-banding and all contained one normal and one deleted chromosome 3. Inconsistent hybridization of FISH probes has been observed previously (personal communication, Dr. Susan Olson) when the probe sequence is adjacent to the breakpoint of a deleted chromosome, possibly reflecting a vulnerability to probe drop off. Southern blot analysis was done to determine if a portion of the cirrin gene was deleted in GM10985. No junctional fragments were observed in GM10985, indicating the cirrin gene is either intact on both copies of chromosome 3 or completely missing on the deleted chromosome 3. However, our FISH results, in which two copies are present, contradicts the latter explanation. It is possible that a deletion would not be observed due to the location of the breakpoint and the enzymes used, but this is unlikely given that two enzymes were used. Unexpectedly, junctional fragments were found in cell line GM10922 suggesting that the chromosomal breakpoint lies within the cirrin gene.

Cloning and sequencing of the unique junctional fragments will allow the chromosomal breakpoint of GM10922 to be delineated.

Our FISH and Southern results have allowed us to more accurately place WI-11041 (cirrin EST) with respect to chromosome 3 markers. Drumheller and coworkers have shown that marker D3S1597 is deleted in four 3p- cell lines including GM10922 and GM10985 (Drumheller, McGillivray et al. 1996). Our FISH analysis with cirrin on cell line GM10985, in combination with Drumheller and coworkers PCR analysis on isolated der(3) chromosomes, indicates placement of WI-11041 (cirrin EST) to be centromeric to marker D3S1597. This placement is in contradiction to the CEPH genetic map, GeneMap, and integrated map at NCBI which all place WI-11041 telomeric to marker D3S1597. The breakpoint of GM10922 was determined to lie between markers D3S1585 and D3S1263 by PCR analysis on an isolated der(3) chromosome (Drumheller, McGillivray et al. 1996). Our Southern blot analysis on GM10922, suggesting the chromosomal breakpoint lies within the cirrin genomic structure, indicates that the cirrin gene should lie between markers D3S1585 and D3S1263 (Figure 15).

Candidate Genes

The critical region for congenital heart defects was originally mapped to a 4 cM region between D3S1250 and D3S18 (Phipps, Latif et al. 1994). Drumheller and coworkers narrowed the critical region to approximately 2 cM between markers D3S1585 and D3S1317 (Drumheller, McGillivray et al. 1996). This region has subsequently been narrowed to a <1000 kb region by Green and coworkers lying somewhere between markers D3S1585 and D3S587 (Green, Latif et al. 1998). Based on size alone, at least 30 or more genes could lie within the CHD critical region. Twenty-eight ESTs and five

genes mapped to the interval between D3S1585 and D3S1317 on the Whitehead Institute and NCBI integrated maps. The genes include the gamma-aminobutyric acid transporter (*GAT1*), a human RNA helicase-like protein (HRH1), the secretory pathway gene *SEC13R*, an uncharacterized gene designated KIAA0121, and the plasma membrane calcium ATPase isoform 2 gene (*PMCA2* or *ATP2B2*). Green and coworkers excluded *PMCA2* as well as another strong candidate, fibulin-2, an extracellular matrix protein (Green, Latif et al. 1998). Recently, mutations in *PMCA2* have been shown to cause deafness and imbalance in mice (Kozel, Friedman et al. 1998; Street, McKee-Johnson et al. 1998). GAT1 has neural specific expression where it functions in the removal of gamma-aminobutyric acid (GABA) from the synaptal cleft (Liu, Lopez-Corcuera et al. 1993) and therefore is not a strong candidate. HRH1 protein facilitates the nuclear export of spliced mRNA by releasing the RNA from the spliceosome (Ohno and Shimura 1996). A defect in HRH1 would have broad phenotypic effects and therefore can be ruled out as a candidate for CHD. Drumheller and coworkers suggest that the *Sec13R* gene is a viable candidate for CHD in 3p- syndrome (Drumheller, McGillivray et al. 1996). The yeast Sec13 is required for vesicle biogenesis from endoplasmic reticulum (ER) during the transport of proteins (Swaroop, Yang-Feng et al. 1994). Again, a defect in a protein involved in vesicle formation for protein transport from the ER to the Golgi apparatus would disrupt transport of a variety of proteins. Finally, KIAA0121 is the designation for the coding sequence of one of the 40 new genes deduced by Nagase and coworkers (Nagase, Seki et al. 1995). If one looks at the GeneMap at NCBI, an additional three genes are found in this region including xeroderma pigmentosum complementation

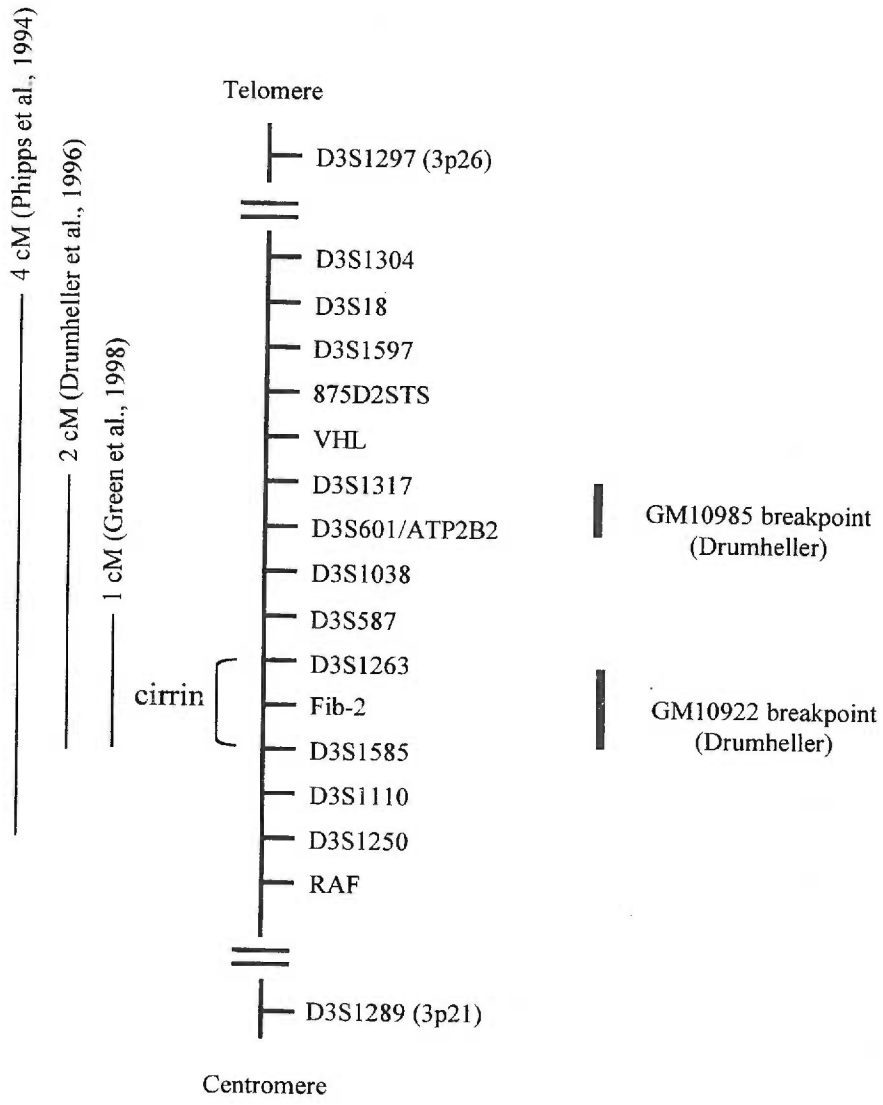


Figure 15. Placement of cirrin (bracket) between markers D3S1585 and D3S1263. Note, cirrin is proximal to D3S1597 in contradiction to the CEPH genetic map, Genemap, and NCBI integrated maps. On the right, breakpoints for GM10985 and GM10922 (determined by Drumheller et al., 1996). On the left are vertical lines representing the narrowing of the critical region for CHD associated with 3p- syndrome. Note, distance between markers is not to scale.

group C (*XPC*), the solute carrier family 6, member 6 (*SLC6A6)* and another uncharacterized gene designated KIAA0763. XPC is involved with the initiation of global genome nucleotide excision repair (Sugasawa, Ng et al. 1998). SLC6A6 is a transporter of taurine, a major intracellular amino acid involved in bile acid conjugation, modulation of calcium flux and neural excitability, osmoregulation, detoxification and membrane stabilization (Uchida, Kwon et al. 1992). KIAA0763 is an unidentified human gene, isolated initially from the brain, with unknown function (Nagase, Ishikawa et al. 1998). Of the genes listed here, none, with the exception of fibulin-2 which has been excluded (Green, Latif et al. 1998), are as strong a candidate as cirrin. The putative extracellular protein has many similarities to known extracellular matrix proteins. Additionally, expression in the endocardial cushions and our FISH results make cirrin a prime candidate.

Identification of homologues

Searches for homologous gene and protein sequence in databases show that cirrin is highly conserved across species. The considerable extended amino acid similarity between cirrin, Chinese hamster HT protein (47% identity, 60% similar) and *C. elegans* F09E8 (44% identity, 58% similar) indicates that these proteins are closely related. However, the HT protein sequence terminates shortly after the carboxy-terminal cbEGF domain, and does not have a predicted transmembrane domain. In addition, the cirrin and HT cDNA sequences are somewhat divergent (42% identical), whereas the human and mouse cirrin cDNA sequences are approximately 85% identical (based on EST sequences). This suggests that HT protein may not be the Chinese hamster homologue of

cirrin, but may instead be a more distantly related member of a cirrin family of genes. By contrast, the predicted amino acid sequence for the *C. elegans* gene F09E8 is overall 58% similar to the human cirrin sequence and includes a predicted carboxy-terminal transmembrane domain. If F09E8 is the *C. elegans* homologue of cirrin, conservation across species from worm to human would suggest that it is functionally important.

Possible phenotypic association

Given the pattern of cirrin expression in early development it would not be surprising to see a more extensive phenotype associated with haploinsufficiency. The most striking physical difference between 3p- syndrome patients with and without a cirrin deletion is the occurrence of a heart malformation. However, Drumheller and co-workers established an additional correlation between the chromosomal breakpoint of the patient samples they analyzed and phenotypic characteristics of the patients. They made the observation that micrognathia and thin upper lips were present in patients with more proximal breakpoints and absent in their patient with the most distal breakpoint (Drumheller, McGillivray et al. 1996). Expression of cirrin in the branchial arches suggests that it is involved in craniofacial development and hence could also be a candidate for the micrognathia and thin lip phenotype in 3p- syndrome patients. Of the three patients whose cell lines were used in the study reported here, patient GM10985 had a thin upper lip, was reported to have micrognathia, and is not deleted for the cirrin gene. Patient GM10922, also with micrognathia and a thin upper lip, has a cirrin gene deletion. The clinical characterization of patient GM07873 was not sufficiently detailed to be informative. However, the fact that both GM10985 and GM10922 have

micrognathia and thin upper lips, while only GM10922 is deleted for cirrin indicates that cirrin is not responsible for micrognathia or thin upper lips. It is possible that cirrin is only involved in the cardiac phenotype observed in 3p- syndrome although it is present in other tissues. As an example, Holt-Oram syndrome is characterized by upper limb malformations and cardiac septation defects and is caused by mutations in the transcription factor *TBX5* (Basson, Bachinsky et al. 1997; Li, Newbury-Ecob et al. 1997). In addition to expression in the heart and forelimb, *TBX5* is expressed in human trachea, lung and thoracic wall (Li, Newbury-Ecob et al. 1997). However, structural defects are not observed in the trachea and lungs of Holt-Oram syndrome. One possible explanation is redundancy of expression and function within the *TBX* gene family in the unaffected tissues. Alternatively, *TBX5* haploinsufficiency on organ morphogenesis, and likewise cirrins, may differ between tissues. Cirrin, like fibronectin, may have a more specialized role in endocardial cushion formation than just a structural component of ECM.

Future directions

In recent years, evidence for genetic heterogeneity in cardiac malformations has grown. With advances in diagnosis and treatment of CHD, a growing number of families have now been reported with cushion defects being transmitted in an autosomal dominant fashion with incomplete penetrance (O'Nuallain, Hall et al. 1977; Emanuel, Somerville et al. 1983; Digilio, Marino et al. 1993). A number of these families have been excluded by linkage to chromosome 21 and/or 8, two sites that are known to contain genes involved with cardiac development (Wilson, Curtis et al. 1993; Cousineau, Lauer et al. 1994; Gennarelli, Novelli et al. 1994; Amati, Mari et al. 1995). Providing further evidence for

genetic heterogeneity, an ECD susceptibility gene was recently mapped to chromosome 1 (Sheffield, Pierpont et al. 1997). Given the nature of CHD in 3p- syndrome, we speculate that disruption of cirrin may also contribute to non-syndromic endocardial cushion defects. Blood, for immortalization of cells and isolation of DNA, is being collected from individuals who have had surgically corrected complete atrioventricular canal defects or atrial septal defects of the ostium primum type. Both FISH and mutation analysis of cirrin will be performed. This patient population was chosen because their cardiovascular defects most closely resemble those seen in 3p- syndrome. Detection of mutations within this population would demonstrate a probable role in cardiac development.

Characterization of the cirrin gene and its protein product has opened many other avenues of investigation. Demonstration of cirrin's involvement in heart development is crucial for proving our hypothesis. To determine cirrin's function, a knock-out mouse to be used for phenotypic studies is being created. This avenue of investigation is underway at present. Additionally, generation of an antibody that recognizes cirrin in a native state is important for determining protein localization as well as for identification of interacting proteins. It was hypothesized that cirrin was involved in proper formation of the endocardial cushions. I further speculated that cirrin may be a component of the cardiac adheron. The use of antisense oligonucleotides or antibodies to cirrin in an *in vitro* collagen gel bioassay, used to examine the role of specific molecules in the epithelial-mesenchymal transformation (Runyan and Markwald 1983; Potts, Dagle et al. 1991; Rezaee, Isokawa et al. 1993; Nakajima, Krug et al. 1994; Sinning and Hewitt

1996), should indicate whether cirrin is involved with endocardial cell transformation or mesenchymal cell migration.

Summary

Cirrin is a highly conserved extracellular protein that may be the founding member of a previously uncharacterized family. Its expression in the endocardial cushions, similarity to extracellular matrix proteins, and its correlating deletion in 3p- individuals with congenital heart defects make cirrin an excellent candidate for heart malformations associated with 3p- syndrome. Its possible role in cardiac morphogenesis and association with endocardial cushion defects in 3p- syndrome leads us to speculate that cirrin may also be involved in the pathogenesis of non-syndromic cardiac septal defects.

Forward Sequencing Primers

Primer		cDNA	Genomic
2-seq	5′ -TCTCCACCTCCCCAGTCTTCTC-3′	94-115	94-115
4-seq	5′ -GGTATGTTCGGCTTGTTTTGGC-3′	627-648	---------
6-seq	5′ -CGTGAACACTGAGGGCTCCTAT-3′	786-807	3614-3635
9-seq	5′ -GTGTTCACCGCCATCTTCATT-3′	1162-1182	4915-4935
11-seq	5′ -CGGGAGAGAACAAGCAGTGTGA-3′	944-965	4160-4181
12-seq	5′ -AAGCCTCTCCACGCCCTCTATC-3′	-(46-25)	-(46-25)
14-seq	5′ -ATTGATGAGTGTGGCACAGAGG-3′	736-757	3564-3585
15-seq	5′ -GTGTGCAGCAAGTCAGACT-3′	289-307	1534-1553
17-seq	5′ -ACTTCGAGTGCCACCGCCTGCTG-3′	305-327	1551-1573
19-seq	5′ -CTCTTCCAGTGGCTGTGCTCA-3′	388-408	1977-1997

Reverse Sequencing Primers

Primer		cDNA	Genomic
1-seq	5′ -ATCCACCTTCAGGGCTACTT-3′	1442-1461	5195-5214
3-seq	5′ -GCCAAAACAAGCCGAACATACC-3′	627-648	---------
5-seq	5′ -CCAGTCCCCGCCAGGTGAAG-3′	1472-1491	5225-5245
7-seq	5′ -GCACAGTCTCGGCACTCATAGG-3′	803-824	---------
8-seq	5′ -AACCAAGCTGTCCTGAGTGTCC-3′	1332-1353	5085-5106
10-seq	5′ -GTGAACACCAAGTCGCCCTTAG-3′	1148-1169	4901-4922
13-seq	5′ -TTGCGTTCTGCCTCAAAGTAGC-3′	593-614	2853-2874
16-seq	5′ -CACACCTTTTGGCTACCT-3′	---------	2900-2917
18-seq	5′ -CCCTGCCCAGTGCCACACCTT-3′	---------	2910-2930
20-seq	5′ -GCAGGAGGGCCCGAAGGTGCC-3′	436-456	2025-2045
MG311	5′ -ACTGCTTGTTCTCTCCCGGACA-3′	940-961	4156-4177

Vector primers

T3 5’-AATTAACCCTCACTAAAGGG-3’

T7 5’-TAATACGACTCACTATAGGG-3’

U19 5’-GTTTTCCCAGTCACGACG-3’

Forward NIRCA Primers

Primer		cDNA	Genomic
11041-1-5′	5′ -GGTAAAAAGTAGCCCTGAAG-3′	1436-1455	5189-5208
11041-2-5′	5′ -GCTTGTAATAACGCAGAT-3′	-(129-112)	-(129-112)
11041-3-5′	5′ -TAGTGTGCCTGGCTTGCT-3′	---------	4672-4689
11041-4-5′	5′ -CCCTTCTCAGGCTTCAGA-3′	---------	3724-3741
11041-5-5′	5′ -GAAGTCCAGCTAGTCTGC-3′	---------	3411-3428
11041-1T7	5′ -GGGCCTGACTCCTTCAGT-3′	-(65-48)	-(65-48)
11041-2T7	5′ -GGAGAAGGGACACGAGGG-3′	511-528	2771-2789
11041-3T7	5′ -GGGAGTTTCTGGGGAGAC-3′	---------	4714-4741
11041-4T7	5′ -GGCCTCCGCTTCTGGAG-3′	---------	3751-3767
11041-5T7	5′ -GGCAAGACCATTCCCCAA-3′	---------	3478-3495
11041-6T7	5′ -TGGGAGGAAGAGAATTTGT-3′	223-241	445-463

Reverse NIRCA Primers

Primer		cDNA	Genomic
11041-1-3′	5′ -GAATAAATAAGTGGCAATGG-3′	1660-1679	5413-5432
11041-2-3′	5′ -ATGTAGAGCCCGGTCTCA-3′	---------	4385-4402
11041-3-3′	5′ -CCTGGGGGTGCTC-3′	---------	3740-3752
11041-4-3′	5′ -CAGCCAGGTTGAGATTTC-3′	---------	3185-3202
11041-1SP6	5′ -GATGGCTGGCGTTGCGTT-3′	607-625	2868-2885
11041-2SP6	5′ -GACCACCTAATCCTAAGC-3′	1558-1575	5311-5328
11041-3SP6	5′ -GGAGCTGCCACTGCTACC-3′	---------	4330-4347
11041-4SP6	5′ -GGAAGGTTTGGAGGGACA-3′	---------	3707-3724
11041-5SP6	5′ -GGATCTGCTCCTTCAC-3′	1030-1045	4246-4261
11041-6SP6	5′ -GGAGATTAGGTAGTTTAG-3′	---------	3158-3175

- Primers with the T7 designation have the following T7 consensus promoter sequence attached at the 5' end:

  5'-TAATAACGACTCACTATAGG(G/A)XXX

- Primers with the SP6 designation have the following SP6 consensus promoter sequence attached at the 5' end:

  5'-ATTTAGGTGACACTATAG(G/A)AXXX

Forward SSCP Primers

Primer		cDNA	Genomic
1F	5′ -CTCTCCACGCCCTCTATC-3′	-(42-25)	-(42-25)
1-2F	5′ -CTGACTCCTTCAGTGAAGCC-3′	-(61-42)	-(61-42)
2F	5′ -ACAAAGACAGTGAGACCC-3′	248-265	---------
2-1F	5′ -CTGGTAGAGGTGCTGGAGGG-3′	268-287	1514-1533
3F	5′ -GGCCTTGGCTACTTTGAG-3′	586-603	2846-2863
3-2F	5′ -TGTGGCCTTGGCTACTT-3′	583-599	2846-2859
4F	5′ -TGGCTTCAGCTTCCCTA-3′	--------	3502-3518
5F	5′ -AAACCTTCCCCTTCTCAG-3′	--------	3716-3734
6F	5′ -GCCCTAGCAGGACTCTG-3′	--------	4094-4110
7F	5′ -AGGAACAGGGATACGAGT-3′	--------	4755-4772
8F	5′ -AAGCCTCTCCACGCCCTCTATC-3′	-(46-25)	-(46-25)
9F	5′ -CGTGGATTTAAGTTTCAT-3′	--------	247-264
10F	5′ -TCTACCGCTAGATTTGAA-3′	--------	1327-1344
10-2F	5′ -TTCCCAGAACCATGACCC-3′	--------	1360-1371
11F	5′ -GAAAGGGCATTGGTCAGAT-3′	--------	1812-1830
12F	5′ -AGAGGGAGAGGGAGAAAA-3′	--------	2643-2660
12-2F	5′ -TGTATAGATGACCTCACC-3′	--------	2682-2699
13F	5′ -GCAGTGTGAAGGAGAA-3′	--------	2761-2776

Reverse SSCP Primers

Primer		cDNA	Genomic
1R	5′ -GTGGCACTCGAAGTCTGA-3′	301-318	1547-1564
1-2R	5′ -AGCAGGCGGTGGCACTCG-3′	309-326	1555-1572
2R	5′ -CCTACCCGAACATACCA-3′	--------	2886-2902
2-1R	5′ -GAGGATGGGCAGGCAGGTGC-3′	--------	2940-2959
3R	5′ -AGGCCCCTGACCATTT-3′	--------	3106-3123
3-2R	5′ -CATGAAGGTGGAGATTAGGT-3′	--------	3165-3185
4R	5′ -GCAGCTCTTCCTCCACTC-3′	--------	3681-3698
5R	5′ -CCAGCCTCTTCACCAT-3′	--------	3976-3992
6R	5′ -TGCCTTCTCTTTGAATGA-3′	--------	4295-4312
7R	5′ -GGGTGGGAGGAGGT-3′	--------	5037-5050
8R	5′ -TGGCACATGCTAGCACTTCAC-3′	--------	338-359
9R	5′ -TGCCTTGTTTACTGCTAT-3′	--------	617-624
10R	5′ -CTCCTGCTGCTTGTG-3′	364-378	---------
10-2R	5′ -GCCTCCTGCTGCTTGTG-3′	364-380	---------
10-3R	5′ -AGGAGCACAACCAGGACA-3′	--------	1714-1731
11R	5′ -GCGAACTAGGGACAGAGC-3′	--------	2112-2129
12R	5′ -GGGGCCAAAACAAGC-3′	--------	2981-2995
13R	5′ -AGAGAATAAAGGACCAAG-3′	--------	3125-3142

Cirrin cDNA Sequence

```
 -365  GGATTGGCCT CGAGGCAAGA TTCGGCACGA GGCTAATTCT GCGGATCCGG CCCCTAATAT
 -305  TCTTTATCAG ACCCTCAGAC AAGAGGCTGA CTTCTGCCCC CTTGTCAAGG AGCGAGGCCA
 -245  CTTTCCTCTC CACCCCATGC TAGCGAGGAT AACTTATTTC TCTTCTGGAA TTGCATCTTA
 -185  TGCGCCTTTC CCCACCCATC CCCACAGCCC CTGCAATACC CAGTTTGGCC TCTTTTGCTT
 -125  GTAATAACGC AGATCCCAGC GCCACGGCAC CTTAGAACAG ACCTTTTTCT TTCTCGCGTG
  -65  GGGCCTGACT CCTTCAGTGA AGCCTCTCCA CGCCCTCTAT CTGCAGGTCC CCAGCCTGGG
   -5  TAAAGATGGC CCCATGGCCC CCGAAGGGCC TAGTCCCAGC TGTGCTCTGG GGCCTCAGCC
   56  TCTTCCTCAA CCTCCCAGGA CCTATCTGGC TCCAGCCCTC TCCACCTCCC CAGTCTTCTC
  116  CCCCGCCTCA GCCCCATCCG TGTCATACCT GCCGGGGACT GGTTGACAGC TTTAACAAGG
  176  GCCTGGAGAG AACCATCCGG GACAACTTTG GAGGTGGAAA CACTGCCTGG GAGGAAGAGA
  236  ATTTGTCCAA ATACAAAGAC AGTGAGACCC GCCTGGTAGA GGTGCTGGAG GGTGTGTGCA
  296  GCAAGTCAGA CTTCGAGTGC CACCGCCTGC TGGAGCTGAG TGAGGAGCTG GTGGAGAGCT
  356  GGTGGTTTCA CAAGCAGCAG GAGGCCCCGG ACCTCTTCCA GTGGCTGTGC TCAGATTCCC
  416  TGAAGCTCTG CTGCCCCGCA GGCACCTTCG GGCCCTCCTG CCTTCCCTGT CCTGGGGGAA
  476  CAGAGAGGCC CTGCGGTGGC TACGGGCAGT GTGAAGGAGA AGGGACACGA GGGGGCAGCG
  536  GGCACTGTGA CTGCCAAGCC GGCTACGGGG GTGAGGCCTG TGGCCAGTGT GGCCTTGGCT
  596  ACTTTGAGGC AGAACGCAAC GCCAGCCATC TGGTATGTTC GGCTTGTTTT GGCCCCTGTG
  656  CCCGATGCTC AGGACCTGAG GAATCAAACT GTTTGCAATG CAAGAAGGGC TGGGCCCTGC
  716  ATCACCTCAA GTGTGTAGAC ATTGATGAGT GTGGCACAGA GGGAGCCAAC TGTGGAGCTG
  776  ACCAATTCTG CGTGAACACT GAGGGCTCCT ATGAGTGCCG AGACTGTGCC AAGGCCTGCC
  836  TAGGCTGCAT GGGGGCAGGG CCAGGTCGCT GTAAGAAGTG TAGCCCTGGC TATCAGCAGG
  896  TGGGCTCCAA GTGTCTCGAT GTGGATGAGT GTGAGACAGA GGTGTGTCCG GGAGAGAACA
  956  AGCAGTGTGA AAACACCGAG GGCGGTTATC GCTGCATCTG TGCCGAGGGC TACAAGCAGA
 1016  TGGAAGGCAT CTGTGTGAAG GAGCAGATCC CAGAGTCAGC AGGCTTCTTC TCAGAGATGA
 1076  CAGAAGACGA GTTGGTGGTG CTGCAGCAGA TGTTCTTTGG CATCATCATC TGTGCACTGG
 1136  CCACGCTGGC TGCTAAGGGC GACTTGGTGT TCACCGCCAT CTTCATTGGG GCTGTGGCGG
 1196  CCATGACTGG CTACTGGTTG TCAGAGCGCA GTGACCGTGT GCTGGAGGGC TTCATCAAGG
 1256  GCAGATAATC GCGGCCACCA CCTGTAGGAC CTCCTCCCAC CCACGCTGCC CCCAGAGCTT
 1316  GGGCTGCCCT CCTGCTGGAC ACTCAGGACA GCTTGGTTTA TTTTTGAGAG TGGGGTAAGC
 1376  ACCCCTACCT GCCTTACAGA GCAGCCCAGG TACCCAGGCC CGGGCAGACA AGGCCCCTGG
 1436  GGTAAAAAGT AGCCCTGAAG GTGGATACCA TGAGCTCTTC ACCTGGCGGG GACTGGCAGG
 1496  CTTCACAATG TGTGAATTTC AAAAGTTTTT CCTTAATGGT GGCTGCTAGA GCTTTGGCCC
 1556  CTGCTTAGGA TTAGGTGGTC CTCACAGGGG TGGGGCCATC ACAGCTCCCT CCTGCCAGCT
 1616  GCATGCTGCC AGTTCCTGTT CTGTGTTCAC CACATCCCCA CACCCCATTG CCACTTATTT
 1676  ATTCATCTCA GGAAATAAAG AAAGGTCTTG GAAAGTT
```

Cirrin Genomic DNA

-365 ggattggcct cgaggcaaga ttcggcacga ggctaattct gcggatccgg cccctaatat
-305 tctttatcag accctcagac aagaggctga cttctgcccc cttgtcaagg agcgaggcca
-245 ctttcctctc caccccatgc tagcgaggat aacttatttc tcttctggaa ttgcatctta
-185 tgcgcctttc cccacccatc cccacagccc ctgcaatacc cagtttggcc tcttttgctt
-125 gtaataacgc agatcccagc gccacggcac cttagaacag acctttttct ttctcgcgtg
-65 gggcctgact ccttcagtga agcctctcca cgccctctat ctgcaggtcc ccagcctggg
-5 taaag**ATGGC CCCATGGCCC CCGAAGGGCC TAGTCCCAGC TGTGCTCTGG GGCCTCAGCC**
56 **TCTTCCTCAA CCTCCCAGGA CCTATCTGGC TCCAGCCCTC TCCACCTCCC CAGTCTTCTC**
116 **CCCCGCCTCA GCCCCATCCG TGTCATACCT GCCGGGGACT GGTTGACAGC TTTAACAAG**g
176 tgggtgcacc ggcagcctcg ttagagggga acacagcgat ttagagtggg gaactctggg
236 atgcaaatct gcgtggattt aagtttcatc ttggtctctt actagttgta tggccctagg
296 caggttgcct ttctgtgcct cagtttccta gtcagtagaa cagtgaagtg ctagcatgtg
356 ccaggcactg tacttagcta ttactaattt tctgtttcca g**GGCCTGGAG AGAACCATCC**
416 **GGGACAACTT TGGAGGTGGA AACACTGCCT GGGAGGAAGA GAATTTGTCC AAATACAAAG**
476 **ACAG**gtaagg ggctgctggg ggaaggggtg tatattcccc tccccgccaa atctctgctc
536 tgctggtgta gggctaggaa ctcttgggga gcacttattc attcaacaaa tagcactgaa
596 acatctatag tatagcagta aacaaggcaa gcaaaatgcc cccttcctgg agctcacatt
656 ctantataaa aaganaagca ntgaatgagt aantgaataa tattatgtcc gatgaaaaac
716 aaacantgaa caccgtaaac tgcangacgt gggaagcaag gtgttcaagg tttggctttg
776 aaccaaaaac ngaattnaaa accaactttt tactggaagg aattggccca agantggact
836 naccttgttt tttttaaaaa ttgcttgaat tgaattaacc cnngggaaaa ttctnaaaag
896 tttcagtgga ggaaattnga ttgtttcaac ntttgggtgt tgngaatagt gnngntatga
956 acattttnaa caagtttttg ttagaaaccn gttttccatt ntngggggtn aatatccagt
1016 aggggaattg gcagttcnta tggcaattcc atgttntant tanggaggaa ttgtttttcc
1076 aacagtggtg atgccatttt gtattcccac ctgcagggng ggaggatttt taatttttca
1136 aacatccctg cnaacanttg ttataggatg tatttttagc catatctacc cagcaagggt
1196 attatattcc attgctcaga tgcaggaact gatgtatgtt acaacaaacc tgcagggaag
1256 gtattgtcat ccccatttta cagatgacaa aacaaagagg ttcagagagg ttaagtgact
1316 tgcccactag atctaccgct agatttgaac ccaggtctct ctgcttccca gaaccatgac
1376 cccttccatt atacctcatg gcctctcctt tgatattttc accgcacgag gaagggtgga
1436 gagagacttg aggagggtgg tgggtggggt ggggcatgtt tcccaccagc cctgccctgt
1496 ccgatcag**TG AGACCCGCCT GGTAGAGGTG CTGGAGGGTG TGTGCAGCAA GTCAGACTTC**
1556 **GAGTGCCACC GCCTGCTGGA GCTGAGTGAG GAGCTGGTGG AGAGCTGGTG GTTTCACAA**g
1616 tgagtggcaa agggccttcc ctggaagtgg gtcacaggtg aggcctggtg ataaggcctg
1676 atttggccga gaagcagggg ggtgcatgct ggggcccatg tcctggttgt gctccttcca
1736 aacccaggtc tgctaagaac ttgccggggg acttgcgctc cactttgagc ctcagtttac
1796 ccttctgcca aatggggaaa gggcattggt cagatggcct tttgggtctt atgtccaagc
1856 tgggttgaat cacagattca ggcatggggg aatgggaaca gcacttatga cactatctca
1916 gcacctcctc cccacctccc tccaccctgc ccctgccatc ag**GCAGCAGG AGGCCCCGGA**
1976 **CCTCTTCCAG TGGCTGTGCT CAGATTCCCT GAAGCTCTGC TGCCCCGCAG GCACCTTCGG**
2036 **GCCCTCCTGC CTTC**gtgagt ttttaagttg ctcttgggga tgggagggga ccaccgagtc
2096 cagggatcca gtcctggctc tgtccctagt tcgctgtgtg aactcaggct actcagataa
2156 acttctctgg acctcagttc ttgcctgcct gacagggctg gggagatggg caaatcagtg
2216 gggaaaggct tggagaaagc acaggggcta gactgagtca tatgcagtat agttatcatc
2276 atgtactaaa gaaatgatag acttcccagc cccttcattt cccatgccag atcggactat
2336 ggtaccttcc tggggagggg cgggtgctga cctggagcct ccatgatcgc catccatcct
2396 cacacacagt cctggcacaa gagctgaact tactaccagc ctcttttaga gcagtctttc
2456 ctgaagctac agataaagag ctagctacaa agatggcaaa ttcagtcttt gtaaaagaaa
2516 atattggctg ggcgcggtgg ctcacacctg taatcccaac actttgggag gctgcggtgg
2576 gtggatcatg aggtcaggag atcgagacca tcgtggaggg gggagggagg gaaggcaagg
2636 gagagggaga gggagaggga gaaaatatta tcttgtatat caaggttgta tagatgacct

2696 cacctggttt ggtgtcttcc cacagCCTGT CCTGGGGGAA CAGAGAGGCC CTGCGGTGGC
2756 TACGGGCAGT GTGAAGGAGA AGGGACACGA GGGGGCAGCG GGCACTGTGA CTGCCAAGCC
2816 GGCTACGGGG GTGAGGCCTG TGGCCAGTGT GGCCTTGGCT ACTTTGAGGC AGAACGCAAC
2876 GCCAGCCATC TGGTATGTTC GGgtaggtag ccaaaaggtg tggcactggg caggggcaaa
2936 tggggcacct gcctgcccat cctcatgctg cccccattcc acccagCTTG TTTTGGCCCC
2996 TGTGCCCGAT GCTCAGGACC TGAGGAATCA AACTGTTTGC AATGCAAGAA GGGCTGGGCC
3056 CTGCATCACC TCAAGTGTGT AGgtaagtgg ggccctagct aggtctggga aaatggtcag
3116 gggcctgggc ttggtccttt attctctcaa cagaagcctg ggctaaacta cctaatctcc
3176 accttcatgg aaatctcaac ctggctggga agctggcatc tctgtgtccc acatgccagc
3236 actcaggaag ggagaaggga aagaaagtgc attgaggagt ccaagcattg ttttaagacc
3296 aagctaacgg tggagctcat cactcctgtt tacatcctgt tggccagaac tcaatcacag
3356 ggacacactt agcttcaaga caggttggca aatgtggtct ctggctgggc agcctgaagt
3416 ccagctagtc tgcttctgtg ttggtagaca gcttgcagtc tctgccatac catttaatcc
3476 caggcaagac cattccccaa cggctctggc ttcagcttcc ctactaaata gggattgaaa
3536 ttctcaccct gctcacctct ctgcagACAT TGATGAGTGT GGCACAGAGG GAGCCAACTG
3596 TGGAGCTGAC CAATTCTGCG TGAACACTGA GGGCTCCTAT GAGTGCCGAG gtcagtgtct
3656 acttctgcag aggaggggac gtgaggagtg gaggaagagc tgctccacac ctgtccctcc
3716 aaaccttccc cttctcaggc ttcagagcac ccccaggcct ccgcttctgg agcgtggctc
3776 ccctgggcct aggtgcacat ctcaccctca tctttctctc ctctctccag ACTGTGCCAA
3836 GGCCTGCCTA GGCTGCATGG GGGCAGGGCC AGGTCGCTGT AAGAAGTGTA GCCCTGGCTA
3896 TCAGCAGGTG GGCTCCAAGT GTCTCGgtga gtctcctgct gatgggacac aggcacctgg
3956 gagtgcctca cccagcatga atggtgaaga ggctggaata tgggcaggtg ggggaaggaa
4016 gggtggaatg ttgcctgggc aagggcagag gggagtgttg agagatggac aagatggagt
4076 cagggtgctg ggtgggggcc cctagcagga ctctgacccc tccctccccc caagATGTGG
4136 ATGAGTGTGA GACAGAGGTG TGTCCGGGAG AGAACAAGCA GTGTGAAAAC ACCGAGGGCG
4196 GTTATCGCTG CATCTGTGCC GAGGGCTACA AGCAGATGGA AGGCATCTGT GTGAAGGAGC
4256 AGATCCCAGg tgagccctgg ggcgggagag gggaggtcct cattcaaaga gaaggcaggc
4316 aagccccttc cccaggtagc agtggcagct ccaggccctg ccccatccct actgccaccc
4376 agcccccctgg aggctgcact gagaccgggc tctacatctg atctccaggt tggctctcag
4436 cagccttata ccttccaggg tacaaaggga atcagacctg gcatcaaatc agtctgcctc
4496 cttaacctgt ttcctcatct atccaatggg accagtgttt gccctggcct gctgaaagct
4556 gtcctaagcc gggggtgtgg tgaagaatgc anggttacct ctctgagcct cactttccca
4616 tttagtgaga cagggatggt aactgcccct tgcagggtgg ttttgagcag taaatgtagt
4676 gtgcctggct tgctgggcag gcctggtggc catgatgatc aggtgtgtgg gagtttctgg
4736 ggagactcca agaactacca ggaacaggga tacgagtgcc aggctgcatc tcttgctcct
4796 ctgcagAGTC AGCAGGCTTC TTCTCAGAGA TGACAGAAGA CGAGTTGGTG GTGCTGCAGC
4856 AGATGTTCTT TGGCATCATC ATCTGTGCAC TGGCCACGCT GGCTGCTAAG GGCGACTTGG
4916 TGTTCACCGC CATCTTCATT GGGGCTGTGG CGGCCATGAC TGGCTACTGG TTGTCAGAGC
4976 GCAGTGACCG TGTGCTGGAG GGCTTCATCA AGGGCAGATA Atcgcggcca ccacctgtag
5036 gacctcctcc cacccacgct gcccccagag cttgggctgc cctcctgctg gacactcagg
5096 acagcttggt ttatttttga gagtggggta agcaccccta cctgccttac agagcagccc
5156 aggtacccag gcccgggcag acaaggcccc tggggtaaaa agtagccctg aaggtggata
5216 ccatgagctc ttcacctggc ggggactggc aggcttcaca atgtgtgaat ttcaaaagtt
5276 tttccttaat ggtggctgct agagctttgg cccctgctta ggattaggtg gtcctcacag
5336 gggtggggcc atcacagctc cctcctgcca gctgcatgct gccagttcct gttctgtgtt
5396 caccacatcc ccacacccca ttgccactta tttattcatc tcaggaaata aagaaaggtc
5456 ttggaaagtt

✦ Indicates additional intronic sequence not shown or known.

Cirrin Amino Acid Sequence

```
MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGL   60
ERTIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWW  120
FHKQQEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGH  180
CDCQAGYGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHH  240
LKCVDIDECGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVG  300
SKCLDVDECETEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTE  360
DELVVLQQMFFGIIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR  420
```

References

Alexander, S. M., K. J. Jackson, et al. (1997). "Spatial and temporal expression of the 72-kDa type IV collagenase (MMP-2) correlates with development and differentiation of valves in the embryonic avian heart." Developmental Dynamics **209**: 261-268.

Amati, F., A. Mari, et al. (1995). "Two pedigrees of autosomal dominant atrioventricular canal defect (AVCD): exclusion from the critical region on 8p." American Journal of Medical Genetics **57**: 483-488.

Argraves, W. S., H. Tran, et al. (1990). "Fibulin is an extracellular matrix and plasma glycoprotein with repeated domain structure." Journal of Cell Biology **111**: 3155-3164.
Artavanis-Tsakonas, S. (1997). "Alagille syndrome- a notch up for the Notch receptor." Nature Genetics **16**: 212-213.

Artavanis-Tskonas, S., K. Matsuno, et al. (1995). "Notch Signaling." Science **268**: 225-232.

Atkinson, J. C., M. Ruhl, et al. (1996). "Collagen VI regulates normal and transformed mesenchymal cell proliferation in vitro." Experimental Cell Research **228**: 283-291.

Aula, P. and H. v. Koskull (1976). "Distribution of spontaneous chromosome breaks in human chromosomes." Human Genetics **32**: 143-148.

Basson, C. T., D. R. Bachinsky, et al. (1997). "Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome." Nature Genetics **15**: 30-35.

Beneck, D., M. J. Suhrland, et al. (1984). "Deletion of the short arm of chromosome 3: a case report with necropsy findings." Journal of Medical Genetics **21**(4): 307-310.

Bhandari, D. G., B. A. Levine, et al. (1986). "H-NMR study of mobility and conformational constraints within the proline-rich N-terminal of the LC1 alkali light chain of skeletal myosin." European Journal of Biochemistry **160**: 349-356.

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

Black, B. L. and E. N. Olson (1999). Control of cardiac development by the MEF2 family of transcription factors. Heart Development. R. P. Harvey and N. Rosenthal. San Diego, Academic Press: 131-142.

Bouchey, D., W. S. Argraves, et al. (1996). "Fibulin-1, vitronectin, and fibronectin expression during avian cardiac valve and septa development." The Anatomical Record **244**: 540-551.

Boyer, A. S., C. P. Erickson, et al. (1999). "Epithelial-mesenchymal transformation in the embryonic heart is mediated through distinct pertussis toxin-sensitive and TGFβ signal transduction mechanisms." Developmental Dynamics **21**: 81-91.

Brown, C. B., A. S. Boyer, et al. (1996). "Antibodies to the type II TGFβ receptor block cell activation and migration during atrioventricular cushion transformation in the heart." Developmental Biology **174**: 248-257.

Brown, C. B., A. S. Boyer, et al. (1999). "Requirement of the Type III TGFβ receptor for endocardial cell transformation in the heart." Science in press.

Brueckner, M., P. D'Eustachio, et al. (1989). "Linkage mapping of a mouse gene, *iv*, that controls left-right asymmetry of the heart and viscera." Proceedings of the National Academy of Science USA **86**: 5035-5038.

Burn, J., P. Brennan, et al. (1998). "Recurrence risks in offspring of sdults with major heart defects: results from first cohort of British collaborative study." The Lancet **351**: 311-316.

Cardy, C. M. and P. A. Handford (1998). "Metal ion dependency of microfibrils supports a rod-like conformation for fibrillin-1 calcium-binding epidermal growth factor-like domains." Journal of Molecular Biology **276**: 855-860.

Carmi, R., J. A. Boughman, et al. (1992). "Endocardial cushion defect: further studies of "isolated" versus "syndromic" occurrence." American Journal of Medical Genetics **43**: 569-575.

Clark, E. B. (1995). Epidemiology of congenital cardiovascular malformations. Moss adn Adam's heart disease in infants, children, and adolescents including the fetus and young adult. G. C. Emmanouilides, H. D. Allen, R. A. Riemenschneider and H. P. Gutgesell. Baltimore, Williams and Wilkins: 60-70.

Cousineau, A. J., R. M. Lauer, et al. (1994). "Linkage analysis of autosomal dominant atrioventricular canal defects: exclusion of chromosome 21." Human Genetics **93**: 103-108.

Crossin, K. L. and S. Hoffman (1991). "Expression of adhesion molecules during the formation and differentiation of the avian endocardial cushion tissue." Developmental Biology **145**: 277-286.

Cruz, M. V. D. L., C. Sanchez-Gomez, et al. (1989). "The primitive cardiac regions in the straight tube heart (stage 9-) and their anatomical expression in the mature heart: an experimental study in the chick embryo." Journal of Anatomy **165**: 121-131.

Davis, C. L. (1924). "The cardiac jelly of the chick embryo." Anatomical Records **27**: 201-202.

Degan, W.G.J., M. Agterbos, et al. (1999). "memA/DRS, a putative mediator of multiprotein complexes, is overexpressed in metastasizing human melanoma cell lines BLM and MV3." Biochimica et Biophysica Acta **1444**: 384-394.

Devriendt, K., G. Matthijs, et al. (1999). "Delineation of the critical deletion region for congenital heart defects, on chromosome 8p23.1." American Journal of Human Genetics **64**: 1119-1126.

Digilio, M. C., B. Marino, et al. (1993). "Risk of congenital heart defects in relatives of patients with atrioventricular canal." American Journal of Disease in Children **147**: 1295-1297.

Drumheller, T., B. C. McGillivray, et al. (1996). "Precise localisation of 3p25 breakpoints in four patients with the 3p- syndrome." Journal of Medical Genetics **33**: 842-847.

Duff, K., R. Williamson, et al. (1990). "Expression of genes encoding two chains of the collagen type VI molecule during human fetal heart development." International Journal of Cardiology **27**: 128-129.

Ehlers, M. R. W. and J. F. Riordan (1991). "Membrane proteins with soluble counterparts: role of proteolysis in the release of transmembrane proteins." Biochemistry **30**(42): 10065-10074.

Eisenberg, L. M. and R. R. Markwald (1995). "Molecular regulation of atrioventricular valvuloseptal morphogenesis." Circulation Research **77**(1): 1-6.

Emanuel, R., J. Somerville, et al. (1983). "Evidence of congenital heart disease in the offspring of parents with atrioventricular defects." British Heart Journal **49**: 144-147.

Engel, J. (1989). "EGF-like domains in extracellular matrix proteins: localized signals for growth and differentiation?" Federation of European Biochemical Societies **251**: 1-7.

Engel, J. (1990). "Domains in proteins and proteoglycans of the extracellular matrix with function in assembly and cellular activities." International Journal of Biological Macromolecules **13**: 147-151.

Farrell, M. J., H. Stadt, et al. (1999). "HIRA, a DiGeorge syndrome candidate gene, is required for cardiac outflow tract septation." Circulation Research **84**: 127-135.

Fishman, M. C. and K. R. Chien (1997). "Fashioning the vertebrate heart: earliest embryonic decisions." Development **124**: 2099-2117.

Frank, G. and A. G. Weeds (1974). "The amino-acid sequence of the alkali light chains of rabbit skeletal-muscle myosin." European Journal of Biochemistry **44**: 317-334.

Freedman, S. J., D. G. Sanford, et al. (1996). "Structure and Function of the epidermal growth factor domain of P-selectin." Biochemistry **35**: 13733-13744.

Funderburg, F. M. and R. R. Markwald (1986). "Conditioning of native substrates by chondroitin sulfate proteoglycans during cardiac mesenchymal cell migration." The Journal of Cell Biology **103**: 2475-2487.

Furie, B. and B. Furie (1988). "The molecular basis of blood coagulation." Cell **53**: 505-518.

Ganoza, M. C. and B. G. Louis (1994). "Potential secondary structure at the translational start domain of eukaryotic and prokaryotic mRNAs." Biochimie **76**: 428-439.

Gennarelli, M., G. Novelli, et al. (1994). "Exclusion of linkage with chromosome 21 in families with recurrence of non-Down's atrioventricular canal." Human Genetics **94**: 708-710.

George, E. L., E. N. Georges-Labouesse, et al. (1993). "Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin." Development **119**: 1079-1091.

Glanville, R. W., R.-Q. Qian, et al. (1994). "Calcium binding, hydroxylation and glycosylation of the precursor epidermal growth factor-like domains of fibrillin-1, the Marfan gene protein." Journal of Biological chemistry **269**: 26630-26634.

Green, E. K., F. Latif, et al. (1998). "Molecular delineation of the 3p- syndrome and detailed mapping of a chromosome 3p25 congenital heart disease gene." American Journal of Human Genetics: A10.

Gulcher, J. R., D. E. Nies, et al. (1989). "An alternatively spliced region of the human hexabrachion contains a novel repeat of potential N-glycosylation sites." Proceedings of the National Academy of Science **86**: 1588-1592.

Handford, P., A. K. Downing, et al. (1995). "The calcium binding properties and molecular aorganization of epidermal growth factor-like domains in human fibrillin-1." The Journal of Biological Chemistry **270**: 6751-6756.

Hartley, D. A., T. Xu, et al. (1987). "The embryonic expression of the Notch locus of Drosophila melanogaster and the implications of point mutations in the extracellular EGF-like domain of the predicted protein." The EMBO Journal **6**: 3407-3417.

Hassold, T., N. Chen, et al. (1980). "A cytogenetic study of 1000 spontaneous abortions." Ann. Human Genetics **44**: 151-177.

Hatsuzawa, K., K. Murakami, et al. (1992). "Molecular and enzymatic properties of furin, a Kex2-like endoprotease involved in precursor cleavage at Arg-X-Lys/Arg-Arg sites." Journal of Biochemistry Tokyo **111**: 296-301.

Higginbottom, M. C., J. T. Mascarello, et al. (1982). "A second patient with partial deletion of the short arm of chromosome 3: karyotype 46,XY,del(3)(p25)." Journal of Medical Genetics **19**: 71-73.

Hoffman, J. I. E. (1995). "Incidence of congenital heart disease: I. post-natal incidence." Pediatric Cardiology **16**: 103-113.

Huang, W. Y., H. H. Q. Heng, et al. (1996). "Assignment of the human GATA4 gene to 8p23.1-p22 using fluoresence in situ hybridisation analysis." Cytogenetics and Cell Genetics **72**: 217-218.

Hurle, J. M., G. T. Kitten, et al. (1994). "Elastic extracellular matrix of the embryonic chick heart: an immunohistological study using laser confocal microscopy." Developmental Dynamics **200**: 321-332.

Isokawa, K., E. L. Krug, et al. (1991). "Identification of ectodermal extracellular matrix proteins that inhibit chondrogenesis of early limb mesoderm." Journal of Cell Biology **115**: 447a.

Isokawa, K., M. Rezaee, et al. (1994). "Identification of transferrin as one of multiple EDTA-extractable extracellular proteins involved in early chick heart morphogenesis." Journal of Cellular Biochemistry **54**: 207-218.

Johnson, M. C., R. M. Payne, et al. (1995). "The genetic basis of paediatric heart disease." Annals of Medicine **27**: 289-300.

Kajii, T., A. Ferrier, et al. (1980). "Anatomic and chromosmal anomalies in 639 spontaneous abortuses." Human Genetics **55**: 87-98.

Kanzaki, T., A. Olafsson, et al. (1990). "TGFβ1 binding protein: a component of the large latent complex of TGFβ1 with multiple repeat sequences." Cell **61**: 1051-1061.

Kimble, J., S. Henderson, et al. (1998). "Notch/LIN-12 signaling: transduction by regulated protein splicing." Trends in Biological Sciences **23**: 353-357.

Kitten, G. T., S. J. Kolker, et al. (1996). "Type VI collagen in the cardiac valves and connective tissue septa during heart development." Brazilian Journal of Medical and Biological Research **29**: 1189-1193.

Kitten, G. T., R. R. Markwald, et al. (1987). "Distribution of basement membrane antigens in cyopreserved early embryonic hearts." The Anatomical Record **217**: 379-390.

Klewer, S. E., S. L. Krob, et al. (1998). "Expression of type VI collagen in the developing mouse heart." Dvevlopmental Dynamics **211**: 248-255.

Koleske, A. J., S. Buratowski, et al. (1992). "A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID." Cell **69**: 883-894.

Kozak, M. (1984). "Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs." Nucleic Acids Research **12**: 857-872.

Kozak, M. (1996). "Interpreting cDNA sequences: some insights from studies on translation." Mammalian Genome **7**: 563-574.

Kozel, P. J., R. A. Friedman, et al. (1998). "Balance and hearing deficits in mice with a null mutation in the gene encoding plasma membrane Ca^{2+}-ATPase isoform 2." The Journal of Biological Chemistry **273**: 18693-18696.

Kramer, H.-H., F. Majewski, et al. (1987). "Malformation patterns in children with congenital heart disease." American Journal of Disease Children **141**: 789-795.

Krug, E. L., R. B. Runyan, et al. (1987). "Extracellular matrix from embryonic myocardium elicits an early morphogenetic event in cardiac endothelial differentiation." Developmental Biology **120**: 348-355.

Kuo, C. T., E. E. Morrisey, et al. (1997). "GATA4 transcription factor is required for ventral morphogenesis and heart formation." Genes and Development **11**: 1048-1060.

Latif, F., F.-M. Duh, et al. (1993). "von Hippel-Lindau syndrome: cloning and identification of the plasma membrane Ca(++)-transporting ATPase isoform 2 gene that resides in the von Hippel-Lindau gene region." Cancer Research **53**: 861-867.

Latif, F., K. Tory, et al. (1993). "Identification of the von Hippel-Lindau disease tumor suppressor gene." Science **260**: 1317-1320.

Lee, J. Y., M. Kambe, et al. (1998). "Cloning and characterization of a novel zinc finger protein that associates with nuclear matrix." DNA and Cell Biology **17**: 849-858.

Levin, M., R. L. Johnson, et al. (1995). "A molecular pathway determining left-right asymmetry
in chick embryogenesis." Cell **82**: 803-814.

Li, L., I. D. Krantz, et al. (1997). "Aligille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1." Nature Genetics **16**: 243-251.

Li, Q. Y., R. A. Newbury-Ecob, et al. (1997). "Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family." Nature Genetics **15**: 21-29.

Lin, A. E., A. H. Herring, et al. (1999). "Cardiovascular malformations: changes in prevalence and birth status, 1972-1990." American Journal of Medical Genetics **84**: 102-110.

Lin, Q., J. Schwarz, et al. (1997). "Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C." Science **276**: 1404-1407.

Little, C. D. and B. J. Rongish (1995). "The extracellular matrix during heart development." Experientia **51**: 873-882.

Liu, Q. R., B. Lopez-Corcuera, et al. (1993). "Molecular characterization of four pharmacologically distinct gamma-aminobutyric acid transporters in mouse brain." Journal of Biological Chemistry **268**: 2106-2112.

Lupas, A., M. Van Dyke, et al. (1991). "Predicting coiled coils from protein sequences." Science **252**: 1162-1164.

Lyons, I., L. M. Parsons, et al. (1995). "Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene *Nkx2-5*." Genes Development **9**: 1654-1666.

Maslen, C. L., G. M. Corson, et al. (1991). "cDNA cloning and partial sequence determination of human fibrillin, a candidate gene for the Marfan Syndrome." Nature **352**: 334-337.

Maurer, P. and E. Hohenester (1997). "Structural and functional aspects of clalcium binding in extracellular matrix proteins." Matrix Biology **15**: 569-580.

Maurer, P., U. Mayer, et al. (1992). "High-affinity and low-affinity calcium binding and stability of the multidomain extracellular 40-kDa basement membrane glycoprotein (BM-40/SPARC/osteonectin)." European Journal of Biochemistry **205**: 233-240.

Mayer, U., R. Nischt, et al. (1993). "A single EGF-like motif of laminin is responsible for high affinity nidogen binding." The EMBO Journal **12**: 1879-1885.

McGuire, P. G. (1990). "Urokinase activity in the developing avian heart." Journal of Cell Biology **111**: 239a.

McGuire, P. G. and S. M. Alexander (1993). "Urokinase production by embryonic endocardial-derived cells: regulation by substrate composition." Developmental Biology **155**: 442-451.

Meno, C., Y. Saijoh, et al. (1996). "Left-right asymmetric expression of the TGFβ-family member *lefty* in mouse embryos." Nature **381**: 151-155.

Merrild, U., S. Berggreen, et al. (1981). "Partial deletion of the short arm of chromosome 3." European Journal of Pediatrics **136**: 211-216.

Miosge, N., T. Sasaki, et al. (1998). "Ultrastructural localization of microfibrillar fibulin-1 and fibulin-2 during heart development indicates a switch in molecular associations." Cellular and Molecular Life Sciences **54**: 606-613.

Mjaatvedt, C. H., E. L. Krug, et al. (1991). "An antiserum (ES1) against a particulate form of extracellular matrix blocks the transition of cardiac endothelium into mesenchyme in culture." Developmental Biology **145**: 219-230.

Mjaatvedt, C. H., R. C. Lepera, et al. (1987). "Myocardial specificity for initiating endothelial-mesenchymal cell transition in embryonic chick heart correlates with a particulate distribution of fibronectin." Development Biology **119**: 59-67.

Mjaatvedt, C. H. and R. R. Markwald (1989). "Induction of an epithelial-mesenchymal transition by an *in vivo* adheron-like complex." Developmental Biology **136**: 118-128.

Mjaatvedt, C. H., H. Yamamura, et al. (1999). Mechanisms of segmentation, septation, and remodeling of the tubular heart: endocardial cushion fate and cardiac looping. Heart Development. R. P. Harvey and N. Rosenthal. San Diego, Acadmic Press: 159-177.

Mochizuki, T., Y. Saijoh, et al. (1998). "Cloning of *inv*, a gene that controls left/right asymmetry and kidney development." Nature **395**: 177-181.

Molkentin, J. D., B. L. Black, et al. (1995). "Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins." Cell **83**: 1125-1136.

Molkentin, J. D., Q. Lin, et al. (1997). "Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis." Genes and Development **11**: 1061-1072.

Morgan, D., L. Turnpenny, et al. (1998). "Inversin, a novel gene in the vertebrate left-right axis pathway, is partially deleted in the *inv* mouse." Nature Genetics **20**: 149-156.

Mowrey, P. N., M. J. Chorney, et al. (1993). "Clinical and molecular analyses of deletion 3p25-pter syndrome." American Journal of Medical Genetics **46**: 623-629.

Nagase, T., K. Ishikawa, et al. (1998). "Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro." DNA Research **5**: 277-286.

Nagase, T., N. Seki, et al. (1995). "Prediction of the coding sequences of unidentified human genes. IV The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1." DNA Research **2**: 167-174.

Nakagawa, T., K. Murakami, et al. (1993). "Identification of an isoform with an extremely large Cys-rich region of PC6, a kKex2-like processing endoprotease." FEBS Letters **327**: 165-171.

Nakajima, Y., E. L. Krug, et al. (1994). "Myocardial regulation of transorming growth factor-β expression by outflow tract endothelium in the early embryonic chick heart." Developmental Biology **165**: 615-626.

Nakajima, Y., K. Miyazono, et al. (1997). "Extracellular fibrillar structure of latent TGFβ binding protein-1: role in TGFβ-dependent endothelial-mesenchymal transformation during endocardial cushion tissue formation in mouse embryonic heart." The Journal of Cell Biology **136**: 193-204.

Oda, T., A. G. Elkahloun, et al. (1997). "Mutations in the human Jagged1 gene are responsible for Alagille syndrome." Nature Genetics **16**: 235-242.

Ohno, M. and Y. Shimura (1996). "A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome." Genes and Development **10**: 997-1007.

Olson, S. B. and R. E. Magenis (1988). Cytogenetic aspects of recurrent pregnancy loss. New York, Thieme Medical Publishers. **6:** 191-202.

O'Nuallain, S., J. G. Hall, et al. (1977). "Autosomal dominant inheritance of endocardial cushion defect." Birth Defects: Original Article Series **13**(3A): 143-147.

Payne, R. M., M. C. Johnson, et al. (1995). "Toward a molecular understanding of congenital heart disease." Circulation **91**: 494-504.

Perris, R., H.-J. Kuo, et al. (1993). "Collagen type VI in neural crest development: distribution in situ and interaction with cells in vitro." Developmental Dynamics **198**: 135-149.

Persson, E., M. Selander, et al. (1989). "Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal growth factor-like domain of bovine factor X." Journal of Biological Chemistry **264**: 16897-16904.

Pfaff, M., M. Aumailley, et al. (1993). "Integrin and arg-gly-asp dependence of cell adhesion to the native and unfolded triple helix of collagen type VI." Developmental Dynamics **198**: 167-176.

Phipps, M. E., F. Latif, et al. (1994). "Molecular genetic analysis of the 3p- syndrome." Human Molecular Genetics **3**(6): 903-908.

Potts, J. D., J. M. Dagle, et al. (1991). "Epithelial-mesenchymal transformation of embryonic cardiac endothelial cells is inhibited by a modified antisense oligodeoxynucleotide to tranforming growth factor β3." Proceedings of the National Academy of Science, USA **88**: 1516-1520.

Potts, J. D. and R. B. Runyan (1989). "Epithelial-mesenchymal cell transformation in the embryonic heart can be mediated, in part, by transforming growth factor β." Developmental Biology **134**: 392-401.

Ramer, J. C., R. L. Ladda, et al. (1989). "Two infants with del(3)(p25pter) and a review of previously reported cases." American Journal of Medical Genetics **33**: 108-112.

Ramsdell, A. F. and R. R. Markwald (1997). "Induction of endocardial cushion tissue in the avian heart is regulated, in part, by TGFβ-3-mediated autocrine signaling." Developmental Biology **188**: 64-74.

Rand, M. D., A. Lindblom, et al. (1997). "Calcium binding to tandem repeats of EGF-like modules. Expression and characterization of the EGF-like modules of human Notch-1 implicated in receptor-ligand interactions." Protein Science **6**: 2059-2071.

Reifen, R. M., R. Gale, et al. (1986). "Partial deletion of the short arm of chromosome 3: further delineation of the 3p25-3pter syndrome." Clinical Genetics **30**: 127-130.

Reinhardt, D. P., D. E. Mechling, et al. (1997). "Calcium determines the shape of fibrillin." The Journal of Biological Chemistry **272**: 7368-7373.

Reinhardt, D. P., R. N. Ono, et al. (1997). "Calcium stabilizes fibrillin-1 against proteolytic degradation." The Journal of Biological Chemistry **272**: 1231-1236.

Ren, R., B. J. Mayer, et al. (1993). "Identification of a ten-amino acid proline-rich SH3 binding site." Science **259**: 1157-1161.

Rezaee, M., K. Isokawa, et al. (1993). "Identification of an extracellular 130-kDa protein involved in early cardiac morphogenesis." The Journal of Biological Chemistry **268**: 14404-14411.

Roebroek, A. J. M., J. W. M. Creemers, et al. (1992). "Cloning and functional expression of Dfurin2, a subtilisin-like proprotein processing enzyme of *Drosophila melanogaster* with multiple repeats of a cysteine motif." The Journal of Biological Chemistry **267**: 17208-17215.

Rongish, B. J., C. J. Drake, et al. (1998). "Identification of the developmental marker, JB3-antigen, as fibrillin-2 and its de nove organization into embryonic microfibrous arrays." Development Dynamics **212**: 461-471.

Rozakis-Adcock, M., R. Fernley, et al. (1993). "The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1." Nature **363**: 83-85.

Runyan, R. B. and R. R. Markwald (1983). "Invasion of mesenchyme into three-dimensional collagen gels: a regional and temporal analysis of interaction in embryonic heart tissue." Developmental Biology **95**: 108-114.

Sagredo, J. M. G., A. Q. Castilla, et al. (1981). "The phenotype of partial monosomy 3 (p25-pter) observed in two unrelated patients." Abstracts of the Symposium of the European Society of Human Genetics, Zurich: 387.

Scambler, P., C. Roberts, et al. (1998). "Hira, a gene from DGS/VCFS region, is required for normal embryogenesis." American Journal of Human Genetics **63(suppl)**: A7.

Scartezzini, P., A. Egeo, et al. (1997). "Cloning a new human gene from chromosome 21q22.3 encoding a glutamic acid-rich protein expressed in heart and skeletal muscle." Human Genetics **99**: 387-392.

Schiaffino, S., B. Dallapiccola, et al. (1999). "Molecular genetics of congenital heart disease." Circulation Research **84**: 247-249.

Schott, J.-J., D. W. Benson, et al. (1998). "Congenital heart disease caused by mutations in the transcription factor NKX2-5." Science **281**: 108-111.

Schubert, D. and M. LaCorbiere (1980). "Role of a 16S glycoprotein complex in cellular adhesion." Proceedings of the National Academy of Science **77**: 4137-4141.

Sheffield, V. C., M. E. Pierpont, et al. (1997). "Identification of a complex congenital heart defect susceptibility locus by using DNA pooling and shared segment analysis." Human Molecular Genetics **6**: 117-121.

Sinning, A. R. and C. C. Hewitt (1996). "Identification of a 283 kDa protein component of the particulate matrix associated with mesenchyme formation." Acta. Anat. **155**: 219-230.

Sinning, A. R., C. H. Hewitt, et al. (1995). "A subset of SBA lectin binding proteins isolated from myocardial conditioned medium transforms cardiac endothelium." Acta. Anat. **154**: 111-119.

Slentz-Kesler, K. A., L. P. Hale, et al. (1998). "Identification and characterization of K12 (SECTM1), a novel human gene that encodes a golgi-associated protein with transmembrane and secreted isoforms." Genomics **47**: 327-340.

Spicer, A. P., M. L. Augustine, et al. (1996). "Molecular cloning and characterization of a putative mouse hyaluronan synthase." The Journal of Biological Chemistry **271**: 23400-23406.

Srivastava, D. (1999). Segmental regulation fo cardiac development by the basic helix-loop-helix transription factors dHAND and eHAND. Heart Development. R. P. Harvey and N. Rosenthal. San Diego, Academic Press: 143-155.

Street, V. A., J. W. McKee-Johnson, et al. (1998). "Mutations in a plasma membrance Ca^{2+}-ATPase gene cause deafness in deafwaddler mice." Nature Genetics **19**: 390-394.

Sugasawa, K., J. M. Ng, et al. (1998). "Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair." Molecular Cell **2**: 223-232.

Supp, D. M., D. P. Witte, et al. (1997). "Mutation of an axonemal dynein affects left-right asymmetry in *inversus viscerum* mice." Nature **389**: 963-966.

Swaroop, A., T. L. Yang-Feng, et al. (1994). "Molecular characterization of a novel human gene, SEC13R, related to the yeast secretory pathway gene SEC13, and mapping to a conserved linkage group on human chromosome 3p24-p25 and mouse chromosome 6." Human Molecular Genetics **3**: 1281-1286.

Tasaka, H., E. L. Krug, et al. (1996). "Origin of the pulmonary venous orifice in the mouse and its relation to the morphogenesis of the sinus venosus, extracardiac mesenchyme (spina vestibuli), and atrium." The Anatomical Record **246**: 107-113.

Tazelaar, J., J. Roberson, et al. (1991). "Mother and son with deletion of 3p25-pter." American Journal of Medical Genetics **39**: 130-132.

Timmers, C., M. A. Whitney, et al. (1996). "Refined mapping of the Fanconi anemia group D complementing gene to a 10cM interval in chromosome 3p25.3." American Journal of Human Genetics **59**: A238.

Tolmie, J. L., P. Batstone, et al. (1986). "Partial deletion of the short arm of chromosome 3." Clinical Genetics **29**: 538-542.

Tran, H., W. J. V. Dusen, et al. (1997). "The self-association and fibronectin-binding sites of fibulin-1 map to calcium binding epidermal growth factor-like domains." The Journal of Biological Chemistry **272**: 22600-22606.

Uchida, S., H. M. Kwon, et al. (1992). "Molecular cloning of the cDN for an MDCK cell Na(+)- and Cl(-)-dependent taurine transporter that is regulated by hypertonicity." Proceedings of the National Academy of Sciences **89**: 8230-8234.

Usheva, A., E. Maldonado, et al. (1992). "Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein." Cell **69**: 871-881.

Verjall, M. and J. D. Nef (1978). "A patient with a partial deletion of the short arm of chromosome 3." American Journal of Disease Children **132**: 43-45.

Williamson, M. P. (1994). "The structure and function of prolin-rich regions in proteins." Biochemical Journal **297**: 249-260.

Wilson, L., A. Curtis, et al. (1993). "A large, dominant pedigree of atrioventricular septal defect (AVSD): exclusion from the Down syndrome critical region on chromosome 21." American Journal of Human Genetics **53**: 1262-1268.

Witt, D. R., B. Biedermann, et al. (1985). "Partial deletion of the short arm of chromosome 3 (3p25-3pter)." Clinical Genetics **27**: 402-407.

Wunsch, A. M., C. D. Little, et al. (1994). "Cardiac endothelial heterogeneity defines valvular development as demonstrated by the diverse expression of JB3, an antigen of the endocardial cushion tissue." Developmental Biology **165**: 585-601.

Yamagishi, H., V. Garg, et al. (1999). "A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects." Science **283**: 1158-1161.

Yamamura, H., M. Zhang, et al. (1997). "A heart segmental defect in the anterior-posterior axis of a transgenic mutant mouse." Developmental Biology **186**: 58-72.

Yokoyama, T., N. G. Copeland, et al. (1993). "Reversal of left-right asymmetry: a situs inversus mutation." Science **260**: 679-682.

Yuan, X., A. K. Downing, et al. (1997). "Solution structure of the transforming growth factor β-binding protein-like module, a domain associated with matrix fibrils." The EMBO Journal **16**: 6659-6666.

Zhang, H., S. D. Apfelroth, et al. (1994). "Structure and expression of fibrillin-2, a novel microfibrillar component preferentially located in elastic matrices." The Journal of Cell Biology **124**: 855-862.