Investigating the biological role of *CRELD1* as a candidate gene for congenital heart defects

By

Gameil Taher Fouad

A Dissertation

Presented to the Department of Molecular and Medical Genetics

And the Oregon Health & Science University

School of Medicine

in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

July 2002

School of Medicine

Oregon Health and Sciences University

Certificate of Approval

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

Gameil Taher Fouad

has been approved

[redacted]
Professor in charge of thesis

[redacted]
Member

[redacted]
Member

[redacted]
Member

[redacted]
Member

[redacted]
Member

Table of Contents

Chapter II: Materials and Methods (continued)

Figures and Tables

Abbreviations

Abbreviation:	
ACN	tACE/Cre/Neo self excision cassette
ASD	atrial septal defect
AV	atrioventricular
AVSD	atrioventricular septal defect
BAC	bacterial artificial chromosome
bp	base pair
BSA	bovine serum albumin
cfu	colony forming unit
cM	centi-Morgan
cR	centi-Ray
DNA-AD	DNA activation domain
DNA-BD	DNA binding domain
DTT	dithiothreitol
EC	E-F hand/calcium binding
ECM	extracellular matrix
EMT	epithelial-mesenchymal transition
ES cell	embryonic stem cell

EST	expressed sequence tag
FS	Follistatin-like
GST	Glutathione-S-Transferase
h	hour
HMA	heteroduplex mobility analysis
HRP	horseradish peroxidase
IPTG	isopropyl β-D-thiogalactoside
kb	kilobase
kD	kilo-Dalton
Mb	megabase
MCS	multiple cloning site
min or (')	minute
MTN	multiple tissue Northern
neo	neomycin gene
OT	outflow tract
PCR	Polymerase Chain Reaction
pfu	plaque-forming unit
QDO	quadruple dropout media (SD/-Ade/-His/-Leu/-Trp)
qs	quantity sufficient
RH	radiation hybrid
s or (")	second
SD	synthetic dropout
SSCP	single strand confirmation polymorphism

SLRP	small leucine-rich proteoglycan
TBST	tris buffered saline + 0.05% tween-20
TK	thymidine kinase
TOF	tetralogy of Fallot
UTR	untranslated region
WE domain	tryptophan and glutamic acid rich region
YAC	Yeast artificial chromosome
YPD	Yeast extract, peptone and dextrose (optimal yeast media)
YPH	Yeast protocols handbook

Where possible, abbreviations are for units as defined by the International System of Units (SI) available at http://physics.nist.gov/cuu/Units/index.html.

Acknowledgements

I find that acknowledging the people to whom I owe a debt of gratitude and who have been supportive throughout this process a difficult and challenging task. This is not because I have any difficulty pinpointing who these special people are; rather, I'm not sure how to convey these feelings with a sufficient depth of sincerity, and it is even difficult to decide whom to acknowledge first, lest anyone feel slighted. I should say for clarity I can't imagine having spent this part of my life without each and every person mentioned herein; each of you have enriched my life and allowed (yes, sometimes *pushed*) me to grow not only as a scientist but also as a person.

Of course, it goes without saying that my mentor, Dr. Cheryl Maslen has played a central role in terms of my progression as a scientist. However, the role Cheryl has played over the course of the last six years is much more. She has been understanding and patient, willing to indulge some rather 'outlandish' ideas I've had along the way, afforded me flexibility in both my thinking and approach to bench work over my course of study all the while showing unending enthusiasm for the project at hand, even when my enthusiasm was waning. She has also become a true friend along the way, offering meaningful insight into the inevitably complex life's that all of us have outside the lab. In fact, had I not chosen to work in the Maslen lab, I would still owe Cheryl a resounding "Thanks!" – because it was she who helped orchestrate and coordinate my unorthodox and hasty arrival in the department after leaving a neurobiology graduate program I began at Cornell. Not only have we broken new ground in terms of genetics, we have also blazed a trail of sorts for those students (like myself) who sometimes make mistakes and count on the good will of others to help us undo these errors. In fact, I should take

the opportunity to thank the entire faculty of Molecular and Medical Genetics here at OHSU, including our Chairman Dr. Robb Moses as well as Drs. Brad Popovich, Susan Olson and Haydeh Payami. This group in particular, in addition to Cheryl, opened the door and welcomed me to OHSU, even if I was over a month late.

Next, I must acknowledge Dr. Susan Hayflick. Like Cheryl, Susan immediately took me in and became a friend and confidant while serving as my Temporary Advisory Committee member during my first year. Susan offered me a practical crash-course in graduate student life, departmental mechanics and a thoughtful yet realistic approach to science. In this respect I consider her a second mentor. I spend considerable time with Susan in discussions that, much like those I had with Cheryl, went well beyond science (I recommend time in *the Chair* to all students). To her I credit much of my success in the program. Dr. Hayflick was also a member of my Thesis Committee, and thus I feel compelled to thank her again for this time-consuming and under-appreciated service to students. Speaking of which, I also thank my entire Thesis Advisory Committee. In addition to Cheryl and Susan, Drs. DeAnne Pillers, Matt Thayer and Rob Glanville, as well as my exam committee chair, Dr. Joseph Weiss (as I am writing this acknowledgment post-exam, there can be no perception of conflict of interest!). Much like Dr. Hayflick, Dr. Pillers was a resource upon which I drew both advice and direction. Again, my gratitude. To Matt Thayer I can credit not only my deep appreciation of focused, hypothesis driven science, but also my confidence that this work goes beyond the minimum threshold of an appropriately rigorous doctoral dissertation. The laboratory of Dr. Rob Glanville and the Maslen lab share a close working relationship and Rob's expertise in the area of biochemistry was invaluable in terms of the protein work

described herein. Additionally, Rob is as quick to share reagents as he is advice. I suspect scientific achievement as a whole would advance considerably more rapidly if his collegial approach and integrity were more widespread. I would be remiss if I did not also thank Drs. Scott Stadler and Manfred Baetcher who, while not on my committee contributed valuable discussions that lead to the successful production of the *Creld1* knockout mouse.

Of course, the OHSU Department of Molecular and Medical Genetics is made of many tremendous people and I'd like to acknowledge them here as well, including the indispensable help of the office staff, in particular Shawna Rinne and Michelle Neuhaus – Shawna for her help in journal club/audio visual and in dealings with student health and Michelle for single handedly navigating the by-laws of the Graduate Council on my behalf leading up to my oral defense and exam.

In my experience, the people who truly make this process what it is on a daily basis are our closest allies: fellow graduate students and post-docs. I must express my admiration as well as gratitude to this unique group of people – mostly because we understand one another perfectly in terms of what it means to go through a program such as this one and, in the case of post-docs, have not yet had the memories fade (too much).

Thanks first to Paul Rupp (upon whose work much of my project is based) not only for being a resource in the lab, but also for being such a great pal outside the lab. Paul was my next door neighbor, my co-worker and fellow microbrew aficionado. Denise Quigley who introduced me to the Vollum Rooftop and was always good for throwing a BBQ. Jae Cho who, along with my housemate Krysta Schliss became a 'second' housemate. Phuoc Tran and Bill Chang who (along with others) introduced me to surfing

and coastal camping. Similarly, my first housemate Dr. Steve McGaraughty who put up with my late night shenanigans as well as introduced me into Canadian culture (especially hockey). Together we taught one another racquetball and spent hours on the court. Dr. Kevin Batille not only kept me familiar with Goose Hollow, but also helped teach me the more practical aspects of genomic cloning and restriction mapping. Dr. Brian Nauert (aka Dr.Taproot) taught me the finer points of immunoprecipitation and protein biochemistry as well as how to play Starcraft which added at least 3 months onto my project. He also fed me often, so how could I not be grateful? I'd also love to thank Shelly Saquet, and though she left OHSU years before I finished my thesis, she was a constant source of support that first year. A special thanks to my pals Shannon Dwinnell and Dave Jacobsen. It was Dave with whom I developed a quick and easy friendship that extended from MMG, to dive bars, the Bagdad, Mt. Hood all the way to Whistler and back. We rode mile after mile on our mountain bikes through Forest Park, shot plenty of pool and probably added a good chunk of our stipends to McMinemans bottom line. Dave and I got into more junk than I dare write. Shannon, though she didn't arrive until my second year, became our partner in crime through many of these jaunts as well as trips to the coast and (at least in my case) endless hours of commiserating, grocery shopping, beer drinking and all around fun. Looking back now, I understand the value of these relationships and experiences and I realize I wouldn't change a thing.

I've already mentioned Cheryl and Paul, but there are few key people in the Maslen lab that deserve their own thanks, primarily Darcie Babcock and Dr. Sue Robinson. Darcie played varied roles from my point of view – not only as a friend but also lab manager who managed to keep rotation students as well as Paul and I largely

(though I'm sure not completely) conscientious in terms of our lab practices. As they say, whoever *they* are, this must be like herding cats. Probably worse. Students couldn't get much done without reagents, space and equipment and Darcie saw to each of these aspects flawlessly. Darcie is quick to laugh and as much as a co-worker, Darcie is a member of that larger circle of people (including Paul, Denise, Dave and others) with whom I spent a good deal of time outside the lab. Thanks. And of course, Sue. Perhaps even more than Paul, to Sue I owe a great deal of support from a scientific point of view. Sue was also quick to sympathize with my gripes about student life – all the while keeping me honest. She also kept me positive when I struggled and realistic in terms of my expectations. Hell, Sue even brought me under her wing (along with Ken) when I met with unexpected delays in my work, even as my family moved out of town in anticipation of both my graduation and our next phase of life. Sue did this in spite of her own particular struggles. Sue, all I can say is I hope I gave as much as I got.

When I ventured off the hill, I found another group of people to whom I owe a debt of gratitude. These people enriched my life in many ways and as I look forward, I suspect I'll count them as life long friends. OHSU and Portland are inextricable in my mind, and without putting too fine a point on it, I'd like to acknowledge Jill, Trisha, and Chantal; Karl and Samantha Kaiser, Kathy Jenkins and the Jenkins family (especially Ernie and Elizabeth), Ramona and all our friends on S.E. Linn Street not to mention Eddie and Livia as well as the especially dedicated and humbling "family" at Loaves and Fishes.

Finally Id love to close this chapter of my life by thanking the people who have really shaped who I am: My own family. It is without question that my father, Dr. M.

Taher Fouad served as the inspiration and motivation to undertake this project from the beginning. From my youth I looked up to him as a complex figure who came to America from Egypt nearly 40 years ago and built his success primarily on two things: his education and unwillingness to quit. This translates into all virtually all he does, and I think, for better or worse, has colored so much of what I do. I think I've learned these lessons well. He gave me that tough kind of support I suspect only a father can. I think each of us want to make our parents happy; Dad, I hope I've lived up. Similarly, I must thank my mother, Karen. Over the course of my life I realize there has been one person who I can count on to lend me a sympathetic and non-judgmental ear and that's my mom. I don't know how many hours we chatted on the phone over the years since I left home at 18. I do know that without these conversations I might have made some very different decisions. My mother has been a constant support throughout. Both of my parents have made many sacrifices on behalf of their kids, and the fact that I've been in school *this* long is proof of this claim. In acknowledgment, Id like to dedicate this work in part to each of them.

In addition to my parents, Id like to acknowledge my brother Rheda and my sister Karima. I know that in their own way, they have become invested in my success in this program and I have felt their love and support. Though we have all chosen different directions in life, you never forget the time you spend in those early formative years and how those experiences shape who we become later in life.

After my first couple of years at OHSU, my family grew by two: my partner and wife Anna-Maria, as well as my stepson, Chase. I'd list all our animals, but as some of them were stolen…ahem… "rescued"… I'd better not. These two special people

literally altered their lives in support of me and my commitment to this program. Chase came all the way from Kansas to Oregon where I think he did as much growing as I did. I'm sure I speak for both of us when I say it was not without its own particular challenges (creating our new family from all the different pieces), and Chase, I know that I've learned as much from having you around as you (I hope) have learned from me. I have said often regarding our relationship that, "…I have not been around a kid since I was one…" and watching Chase grow over the last few years has been a gift, not to mention an excuse to behave more like a kid myself (some would argue this is nothing new).

As I reflect on the list above, I realize that the deepest and perhaps most complex relationship I developed over the years is that which I share now with Anna-Maria. She literally gave up a world she had built for herself to come share life with that of a struggling student. I told her I had no money, worked long hours and it was not a very glamorous life. She came anyway. This took tremendous courage and I've often wondered were the circumstances reversed if I would have been so bold. In retrospect, I can see that she offered me both unbelievable patience and every bit of support she possibly could. For both these gifts, I'm as grateful as I am for the gift of her companionship. In addition to my parents, I also dedicate this work to Anna-Maria.

Abstract

Congenital Heart Defects (CHD) are the most frequent form of birth defect, occurring in nearly one percent of newborns. CHD are the leading cause of premature death in infants displaying congenital malformation. Of all congenital heart abnormalities, most often found are those involving improper formation of valves and septa. 3p- syndrome is a rare cytogenetic disorder marked by heterogeneous deletions at the distal end of chromosome 3p. The phenotype is variable and atrioventricular septal defects (AVSD) are observed in individuals with more extensive deletions most likely resulting from haploinsufficiency of one or more deleted genes (AVSD2;OMIM#606217). Recently, the critical region for CHD associated with 3p- syndrome has been narrowed to chromosome 3p25 and *CRELD1* (Cysteine Rich with EGF-Like Domains; formerly cirrin) maps to this area. *CRELD1* encodes a highly conserved protein anchored to the cell membrane that is widely expressed during embryogenesis. Whole mount *in situ* hybridization experiments in chick embryos demonstrate expression of *CRELD1* in the limb buds and branchial arches. CRELD1 is also expressed in the myocardium and endocardial cushion tissue of the developing chick heart. Human multiple tissue northern (MTN) blots analyzed with a *CRELD1*- specific probe show particularly strong expression in heart. A cardiac-specific MTN shows robust *CRELD1* expression throughout the heart, including outflow tract, ventricles and atria. Fetal expression was also detected. Taken together these data support the hypothesis that haploinsufficiency of *CRELD1* results in CHD. In order to investigate the function of CRELD1, two avenues were pursued. The first approach was to test the hypothesis that CRELD1 haploinsufficiency causes CHD by recapitulating the specific subset of CHD found in 3p- patients through the generation of a *CRELD1*$^{+/-}$

mouse. The second avenue was to conduct a yeast two-hybrid interaction screen with the aim of identifying proteins with which CRELD1 interacts. This report details the production of the *CRELD1* knockout mouse. Also presented are the results of a yeast two-hybrid screen which show that *CRELD1* interacts with the small multifunctional extracellular matrix proteins SPARC/Osteonectin/BM-40 and decorin.

Chapter I

Introduction

Introduction

Congenital heart disease: Incidence

Congenital cardiac malformations are among the most common form of birth defect in humans, occurring in nearly one percent of newborns. It has been noted that, excluding pathogenic infection, congenital heart defects (CHD) are the single greatest factor leading to infant mortality (Srivastava and Olson, 2000). In recent decades, techniques allowing successful surgical correction in infants and young children affected with congenital heart malformations have put a premium on efforts to establish the rate with which these anomalies occur. Additionally, the ability to generate animal models of CHD, coupled with a growing appreciation for the underlying genetic programs responsible for embryonic heart assembly have opened new avenues of investigation, with the long term goal of developing novel methods of therapeutic intervention.

The reported incidence of CHD in industrialized nations varies from approximately 4 to 12 per 1000 live births (Hoffman, 1995a). This variation highlights the difficulty with which cardiac malformation is detected. In a particularly interesting study, Grech and Gatt report a frequency of 8.8 per 1000 live births (Grech and Gatt, 1999). Their four-year survey focused strictly on the country of Malta, which has several features making it amenable for this type of epidemiological study. Officially a Roman-Catholic state with a population nearing 400,000 people, the Maltese adhere to a strict policy forbidding termination of pregnancy, yet have aggressive post-natal screening for syndromic congenital cardiac malformation. Whenever a chromosomal anomaly is suspected, or when an infant presents with multiple major abnormalities, echocardiographic assessment is made in order to both detect and formulate appropriate

treatments for children with a heart malformation, taking into account their specific clinical findings. However, due to the nature of this screening program, it is likely that non-syndromic CHD is often overlooked, and thus this number represents at best a conservative figure. Therefore, though consensus estimates of the incidence of CHD are reported to be approximately 1% of live births, the actual overall incidence is probably much higher and it is likely that the majority of affected fetuses expire before term (Hoffman, 1995b).

Cardiovascular development

Assembly of a fully functional cardiovascular system in vertebrates is the result of a complex cascade of events involving cell migration, epithelium formation, autocrine and paracrine signaling, and regional differentiation into distinct cell types. Detailed anatomical observation and experimental manipulation have provided insight into the fluid process of cardiac morphogenesis (Manner, 2000).

The earliest developmental decisions in all vertebrate hearts share a striking similarity (Fishman and Chien, 1997). Studies in avian embryos, most notably chicken and quail, have been of particular utility in elucidating the process of cardiogenesis. In the developing chick embryo two clusters of mesodermal cells migrate along the ventral midline towards the apex of the primitive streak (Hensen's node). A concentration gradient of fibronectin decreasing caudally appears to be critical for this migration, in that antibodies against fibronectin but not other extracellular matrix molecules (i.e. laminin, collagen types I and IV) can disrupt this migration (Linask and Lash, 1988b; Linask and Lash, 1988a). These cells, arranged bi-laterally with respect to the embryonic axis

constitute the presumptive heart forming fields (cardiogenic mesoderm) and are derived from splanchnic mesoderm (Gilbert, 2000). Fate mapping in avian embryos supports the notion that, even at the earliest stages of development, there is a tendency for more caudally located precursors to contribute to more caudally derived portions of the heart tube, and more rostrally positioned structures to be derived from more rostrally positioned populations (Garcia-Martinez and Schoenwolf, 1993).

The migrating splanchnic mesoderm begins to coalesce anteriorly in a crescent shaped epithelial sheet that will later fold and merge to become the heart tube. As would be expected in any tissue forming an epithelium, cell adhesion molecules play an indispensable role. As these cells fuse into the heart tube, expression of N-cadherin, a calcium-dependent cell adhesion molecule, becomes upregulated and cardiac precursors begin to differentiate. In chick embryos incubated with a function-disrupting N-cadherin monoclonal antibody (NCD-2) at 18 hours post-fertilization, hearts fail to form (Linask et al., 1997). However, in N-cadherin-null mice a thin weakly contracting heart is found, suggesting the participation of compensatory cell adhesion molecules in the process of mammalian cardiogenesis (Radice et al., 1997). If the application of NCD-2 in chick embryos is postponed an additional 2 to 8 hours, only the most anterior portions of the differentiating heart tube showed staining with MF20, a monoclonal antibody specific for sarcomeric myosin and an early indicator of cardiac differentiation. Interestingly, the application of NCD-2 beyond the 29th hour of development has no effect on normal heart formation, indicating that N-cadherin mediated adhesion (and perhaps, signaling) has been completed by this time point in avian embryos (Linask et al., 1997).

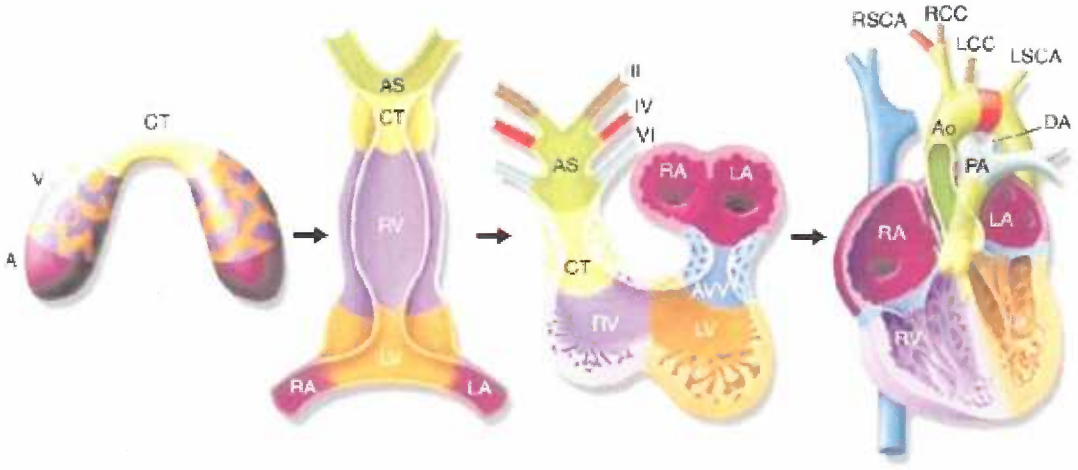
Even as the cardiac epithelium is forming, with obligatory up-regulation of N-cadherin and other cell adhesion molecules, a subpopulation (approximately 5%) of the presumptive cardiac mesoderm is decreasing expression of N-cadherin and is segregating from the still-forming epithelium. Other proteins involved in cell-cell attachment, including neural cell adhesion molecule (N-CAM) are down regulated simultaneously (Mjaatvedt and Markwald, 1989; Crossin and Hoffman, 1991). Studies in quail embryos demonstrate these cells delaminate from the epithelium and migrate inward, later merging to form an endothelial layer expressing the QH1 marker, a hallmark of cells destined to become cardiac endothelium (Coffin and Poole, 1988; Garcia-Martinez and Schoenwolf, 1993). Most cardiac epithelial cells will make up the myocardium; only those QH1 positive cells will later form the endocardium. At this early stage then, the tissues of the myocardium and endocardium are specified as the two-layered heart tube primordium extends. Using a replication defective virus containing a β-gal reporter to perform fate analysis on individual progenitor cells within the heart forming field, it was established that clonal populations of either myocardial or endocardial cells were observed, and in no instance were clonal populations found to contribute to both layers simultaneously (Cohen-Gould and Mikawa, 1996). Thus, it appears specification of a myocardial or endocardial cell fate is made even before the heart forming fields merge to form the primordial heart.

Morphologically, five primitive regions of the extending heart tube become visible under the microscope as the fusion of the cardiogenic mesoderm progresses. The first regions to become visible in the linear heart tube are the primitive right and left ventricles, followed by the atrioventricular segment and sinoatrial segment caudally.

Finally, a fifth primitive region termed the outflow tract (or conus) forms anteriorly of the ventricular primordia (Mjaatvedt, 1999). Historically, it was presumed that each of these morphologically distinct zones represented 'primitive cardiac cavities' which later gave rise directly to each of the mature cardiac cavities, specifically the two atria, two ventricles and the great arteries (Davis, 1927), (de la Cruz, 1989). Using *in vivo* labeling techniques coupled with scanning electron microscopy and standard histology, de la Cruz and others were able to show definitively that the notion of 'primitive cardiac cavities' was inaccurate and proposed instead the idea of 'primitive cardiac regions' (de la Cruz, 1989). Moreover, their work showed that the early heart tube (preceding morphology of the atrioventricular and sinoatrial segments) contains primarily two regions and that these regions are the primordia of only the trabeculated portion of the right and left ventricles, and that each region itself did not anatomically correspond to the complete right or left ventricular chamber (de la Cruz, 1989). It is now widely accepted that the definitive mature heart forms only by the integration and remodeling of the primitive cardiac regions (Mjaatvedt, 1999). A simplified view of human heart morphogenesis, from cardiac crescent to a four-chambered heart is shown in figure 1.

Figure 1

Schematic of cardiac morphogenesis. Illustrations depict cardiac development with color coding of morphologically related regions, seen from a ventral view. Cardiogenic precursors form a crescent (left-most panel) that is specified to form specific segments of the linear heart tube, which is patterned along the anterior–posterior axis to form the various regions and chambers of the looped and mature heart. Each cardiac chamber balloons out from the outer curvature of the looped heart tube in a segmental fashion. Neural crest cells populate the bilaterally symmetrical aortic arch arteries (III, IV and VI) and aortic sac (AS) that together contribute to specific segments of the mature aortic arch, also color coded. Mesenchymal cells form the cardiac valves from the conotruncal (CT) and atrioventricular valve (AVV) segments. Corresponding days of human embryonic development are indicated. A, atrium; Ao, aorta; DA, ductus arteriosus; LA, left atrium; LCC, left common carotid; LSCA, left subclavian artery; LV, left ventricle; PA, pulmonary artery; RA, right atrium; RCC, right common carotid; RSCA, right subclavian artery; RV, right ventricle; V, ventricle.


CT
V
A
AS
CT
RV
RA
LA
II
IV
VI
AS
RA
LA
CT
AVV
ITV
LV
RSCA
RCC
LCC
LSCA
Ao
DA
PA
RA
LA
RV
Day 15
Day 21
Day 28
Day 50

Nevertheless, while the segments of the linear heart tube do not directly correlate with the fully formed and functional chambers of the mature heart, each segment does constitute a distinct region with a specific developmental fate and much like the myocardium and endocardium, begins to express a unique protein profile before the differences can be appreciated morphologically. For example, it was discovered that in avian embryos, ventricular myosin heavy chain-1 (*VMHC1*) was expressed in all myogenic populations during the initial phases of cellular differentiation, but becomes restricted to the presumptive ventricular cardiogenic mesoderm prior to the formation of definitive heart structures (Bisaha and Bader, 1991). Similarly, using RNA isolated from differentiating hearts of chick embryos ranging in age from embryonic day 7 to hatching (21 days), the expression of atrial myosin heavy chain-1 (*AMHC1*) was analyzed. Surprisingly, the results of these experiments showed that AMHC1 is most easily detected in more immature embryos and diminishes during development. Nevertheless, AMHC1 was found to be atrial-specific and was never detected in non-cardiac tissues, including skeletal muscle and brain (Yutzey et al., 1994). Furthermore, *in situ* hybridization experiments showed *AMHC1* lags *VMHC1* expression and rather than being widely expressed in myogenic cells, is restricted only to the posterior portion of the heart tube – that is, the presumptive atrial tissue.

Expression of ventricular myosin light chain-2 (*Mlc2v*) during murine embryogenesis has been examined using a transgenic mouse. This mouse harbored a 250 bp promotor specific for *Mlc2v* fused to a luciferase reporter. Using both measurement of luciferase activity and immunohistochemistry with antibodies directed against luciferase, it was found that by embryonic day 8.5, Mlc2v is restricted to the presumptive

ventricles and, to a lesser degree, the proximal outflow tract of the heart tube. However, just after embryonic day 10, *Mlc2v* expression localized to the ventricular region only. This ventricle-restricted expression of *Mlc2v* precedes chamber septation, but occurs only after formation of the heart tube (O'Brien et al., 1993).

Using a combination of Northern blot analysis, RT-PCR, immunoblotting, *in situ* hybridization and RNAse protection assays, a preferential pattern of expression of myosin light chain 2a (*Mlc2a*) in the presumptive atria was discovered during murine cardiac development (Kubalak et al., 1994). In this study, Mlc2a was not detected in ventricle, skeletal muscle, liver or uterus. Furthermore, *Mlc2a* expression was clearly detectable in differentiating ES cells by day 6 *in vitro*, but not in fibroblast feeder layer cells, which suggests that activation of *Mlc2a* gene expression happens relatively early in cardiogenesis, again preceding distinct morphological differences (Kubalak et al., 1994).

Tissue-specific protein expression is not limited only to those molecules that play a predominant role in the cardiac musculature. Given that the rhythmic contraction of the presumptive ventricles and atria begins long before any definitive conduction system is recognizable, investigators postulated regional differences in cardiomyocytes being responsible for mediating (either by retarding or increasing) the velocity of contractual impulses (Lamers et al., 1987). The finding that acetylcholinesterase is expressed transiently in the cardiomyocytes during the first 24-36 hours of avian embryonic development suggested that a cholinergic system may modulate these contractions. This idea was advanced with the observation that acetylcholinesterase, while initially expressed throughout the myocardium, becomes restricted to the outflow tract over the next 24 hours. The presumptive ventricles, which show intense acetylcholinesterase

activity early on, demonstrate a dramatic decrease in activity over this same interval (Lamers et al., 1990). The implicit conclusion is that a sub-population of embryonic cardiomyocytes has begun to differentiate into what will ultimately be the cardiac conductive tissue, concomitant with the chamber-specific expression of different cardiac myosins (Lamers et al., 1990).

As development proceeds, a population of neural-crest cells, distinct from the cardiogenic mesoderm, migrates caudally from the cranial region towards the outflow tract. These cells are crucial for the structural development of the aortic arch and its derivatives as well as outflow tract septation. Neural crest cell-ablation experiments performed in chick resulted in malformed great arteries and persistent truncus arteriosus (Kirby and Waldo, 1995). More exhaustive route-mapping of migrating neural-crest in quail-to-chick chimeras much later in development showed that, in addition to playing a role in the septation of outflow tract, these cells were also found at the base of the ascending aorta and pulmonary trunk, the semilunar valve leaflets as well as at the site of closure of the ventricular septum (Waldo et al., 1998). These results demonstrate clearly an essential role for neural-crest in cardiovascular development.

At the onset of neural-crest migration into the pharyngeal arches and outflow tract, the incomplete linear heart tube has already initiated the process of rightward looping and spontaneous contractions have begun (33-36 h in chick) (Fishman and Chien, 1997; Farrell et al., 1999). This step is the first in a series of gross morphological movements and tissue remodeling that will transform the immature tubular heart to a fully functional four-chambered organ (Manner, 2000). As mentioned previously, the heart tube (completed only after looping has begun) has five segments: the outflow tract,

the primitive right and left ventricles, the atrioventricular canal and sinoatrial segment. Furthermore, the linear heart consists of two concentric epithelia (the outer myocardium and inner endocardium). Interestingly, even at this early stage, the cells composing the endocardium are not homogenous. A subpopulation of cells has been identified that express the $JB3^+$ marker. A JB3 monoclonal antibody has been produced which recognizes a 350 kD band on Western blots. This marker is thought to be fibrillin-2 and distinguishes those endocardial cells capable of differentiating into endocardial cushion mesenchyme from myocardial precursors (Mjaatvedt, 1999). Between the myocardium and the endocardium is a thickened acellular assembly of extracellular matrix proteins classically called the 'cardiac jelly' (Davis, 1924; Bouchey et al., 1996). This matrix is known to include fibulin-1, vitronectin, fibronectin, flectin, tenascin (cytoactin), CTB proteoglycan as well as other unidentified proteins (Crossin and Hoffman, 1991; Spence et al., 1992; Bouchey et al., 1996; Tsuda et al., 1996). These extracellular proteins are not distributed evenly down the longitudinal axis of the cylindrical heart. Rather, matrix deposition occurs nearly exclusively at both the atrioventricular region (AV) and outflow tract (OT) (Mjaatvedt, 1999). As the process of looping continues, inductive signals from the AV myocardium precipitate delamination of AV endocardial epithelial cells that then transform to mesenchyme and migrate into the extracellular matrix (Runyan and Markwald, 1983; Mjaatvedt and Markwald, 1989). In contrast to AV and OT endothelium, cells from ventricular endothelium are unresponsive to these inductive myocardial signals (Krug et al., 1987).

The mechanisms responsible for the endocardial epithelial-mesenchymal transition (EMT) have been investigated using an *in vitro* assay. This assay involves co-

culturing endothelial cells on a hydrated collagen matrix with extracts presumed to include the signaling factors. When the appropriate factors are present, cultured endothelial cells undergo stereotypical differentiation and subsequent migration into the underlying matrix, resulting in the formation of mesenchymal valve and septal precursors (Krug et al., 1987). Using this approach, intact pre-EMT myocardium, pre-EMT myocardial conditioned media or extracts of embryonic hearts prepared by EDTA treatment were tested for their ability to induce EMT. Krug and co-workers were able to mimic the EMT seen *in vivo* with each of these components, suggesting the atrioventricular myocardium produces a soluble signal that initiates EMT in a paracrine fashion. Further, in the absence of myocardial signals, the endothelial monolayer degenerates after a period of several days, indicating the myocardium produces signals critical for not only for cellular differentiation but also maintenance. In the same study, either EGF or FGF alone were unable to initiate EMT (Krug et al., 1987).

The EDTA-soluble extracts of embryonic chick hearts were found to contain multi-protein aggregates containing fibronectin as well as several lower molecular weight proteins (Mjaatvedt et al., 1987). Purified fibronectin alone was unable to mimic complete extract, though intact isolated aggregates (called 'cardiac adherons') were found to preserve biological activity (Mjaatvedt and Markwald, 1989). Furthermore, anti-serum prepared against isolated EDTA-soluble aggregates was able to block biological activity *in vitro* in a dose-dependent manner. Additionally, this anti-serum, termed ES1 (EDTA-soluble) was able to react with two major and three minor proteins on immunoblot, confirming the multi-protein nature of the adherons (Mjaatvedt et al., 1991). Interestingly, the presence of adherons may have a more general role to play in

development as ES1 antiserum was found to interact with sites of neural-crest formation and migration in limb ectoderm.

Using ES1 antiserum to screen an embryonic cardiocyte cDNA expression library, an extracellular protein named ES/130 was identified (Rezaee et al., 1993). RT-PCR results indicated the ES/130 message was detected in the atrioventricular canal and outflow tract, both regions known to undergo EMT. Monoclonal antibodies prepared against the isolated peptide identified a 130 kD protein from both cultured cardiocytes and conditioned media capable of initiating EMT. Further, these antibodies were capable of blocking EMT, indicating ES/130 was a biologically active component of the cardiac adherons (Rezaee et al., 1993). In 1997, following the identification and isolation of hLAMP-1 (a 283 kD component of the cardiac adherons) Sinning proposed a model of the adherons that placed fibronectin at the core of a multi-subunit protein complex with ES/130, hLAMP-1, and other ECM proteins associated by what is most likely a combination of ionic interaction and disulfide bond formation. In this same study, several additional proteins were isolated as part of the particulate matrix, however many of these components have yet to be identified (Sinning, 1997).

While EGF and FGF were shown to be insufficient in initiating EMT, it has been shown that modified oligonucleotides designed to specifically interfere with TGFβ3 production can block EMT in the *in vitro* collagen gel bioassay (Potts et al., 1991). Interestingly, these oligonucleotides specifically block mesenchymal-cell invasion into the collagen gel, but do not inhibit activation of endothelial cells (characterized by cellular hypertrophy and loss of cell-cell adhesion). Thus the authors conclude that EMT is a two-step process: the first involving endothelial cell activation and the second

marked by cell transformation and migration, and that TGFβ3 mediates the second stage of this process, most likely after the cardiac endothelium receives EMT stimulation from the myocardium. This study did not, however, address whether TGFβ3 was produced by the endocardium, the myocardium or both.

In a follow-up study, Nakajima and co-workers were able to show a second polyclonal antibody to the EDTA-soluble extract of chick hearts (ES3) was able to inhibit TGFβ expression in transforming atrioventricular and outflow tract cardiac endothelium (Nakajima et al., 1994). They proposed inductive signals from the myocardium not only stimulate EMT in AV and OT competent cells, but also regulate endocardial production of TGFβ in a paracrine fashion (Nakajima et al., 1994).

Later experiments from the same laboratory were able to show that TGFβ3 antisense oligos were able to block mesenchyme formation despite the presence of conditioned myocardial medium, and the subsequent addition of recombinant TGFβ3 was able to reverse this effect in explant cultures (Nakajima et al., 1998). From this result, these investigators concluded TGFβ3 is absent in the conditioned myocardial medium. Further study was able to show that production of TGFβ3 by cardiac endothelium competent to undergo EMT is necessary for transformation to mesenchyme. They propose that TGFβ3 produced by endocardium plays an autocrine role by mediating the earliest steps of EMT in the endocardial epithelium (Nakajima et al., 1998).

As mentioned earlier, the expanding pockets of mesenchyme populate the thickened cardiac jelly (both in the atrioventricular canal and outflow tract), specifically in those areas that will ultimately become the endocardial cushions. The main source of this mesenchyme is the migrating endocardial cells following stimulation to undergo

EMT from the AV canal and OT. It is primarily this endocardial cushion tissue, with a contribution from migrating cardiac neural-crest that will eventually give rise to the fully formed septa and valve leaflets of the mature heart. Therefore, perturbations of this process, whether insufficient signaling, defective cell-cell de-adhesion and/or migration, lack of cellular hypertrophy and differentiation or inadequate proliferation would be predicted to result in atrioventricular septal defects (AVSD). It is worth noting that as a class, developmental aberrations of valves and septa comprise the largest subset of congenital heart defects (Eisenberg and Markwald, 1995; Olson and Srivastava, 1996).

Genetics of heart development

Genetic dissection of discrete steps in heart development has led to the emergence of the notion of genetic 'modules' or 'programs' of cardiogenesis with each unique program existing in a spatially and temporally restricted space (Olson and Srivastava, 1996; Fishman and Olson, 1997; Srivastava and Olson, 2000). These ideas have arisen largely through the study of model organisms resulting from work that took advantage of naturally occurring mutations, random mutagenesis or targeted gene deletion. Despite the fact that there now exist numerous examples of genetic defects resulting in heart malformation in such model systems, finding the genetic causes of human CHD has proved more difficult. In fact, though the clinical spectrum of many congenital human heart abnormalities has been well characterized, there remains a paucity of information with regard to the specific underlying genetic defect(s). In part, this is likely due to the high rate of embryonic lethality *in utero* of fetuses harboring the most severe malformations. Further, as mentioned, the accurate detection and diagnosis of CHD at

birth remains problematic. Finally, there is accumulating evidence that many forms of CHD are complex and result from defects in more than a single gene (Novelli et al., 1999; Epstein, 2001).

Despite these difficulties, a handful of the underlying genetic lesions in families affected with syndromic or non-syndromic CHD have been identified, with most of the progress having been made in the last decade. Furthermore, some of these genes have been identified as being involved in normal cardiogenesis on the basis of defects first identified in model systems, followed by a candidate gene approach to scan human orthologs or homologs for mutations. Increasingly, it is appreciated that not only are defects in septation (aortic, atrial or ventricular) the most common class of CHD, but in families with a known mutation, CHD is predominantly a disease of haploinsufficiency marked by both incomplete penetrance and variable phenotype (Olson and Srivastava, 1996; Srivastava and Olson, 2000).

One of the earliest examples of using model systems to aid in clarifying the genetic basis behind human CHD began with the observation that *Drosophila melanogaster* embryos harboring mutations in the homeobox-containing transcription factor *tinman* failed to form a heart (dorsal vessel) as well as visceral mesoderm (Bodmer, 1993). The *tinman* homologue *Nkx2-5* was found to be expressed both in precardiac mesoderm and myocardium of embryonic mice. Targeted disruption of *Nkx2-5* proved lethal during embryogenesis, with heart development arresting in the mutant mouse prior to cardiac looping (e8.5) (Lyons et al., 1995; Harvey, 1996). It was also reported that isolated clusters of cultured ES cells mutant for *Nkx2-5* form cardiac foci and initiate beating. This suggests Nkx2-5 either acts at a different point in heart

development in mice as opposed to flies, or that there exists partial functional redundancy of this factor in mammals. A more general explanation between the differences observed between fly and mouse mutants is that vertebrate cardiogenesis most likely requires additional genetic pathways operating in concert (Harvey, 1996).

Nevertheless, the clear role of *tinman/Nkx2-5* in heart formation captured the attention of investigators studying autosomal dominant non-syndromic ASD and AV conduction delay in four separate kindreds (Schott et al., 1998). This group established linkage in these families to chromosome 5q36 (LOD = 3.91; $\theta = 0$). Since the gene encoding the human homeobox transcription factor NKX2-5 had been mapped to this same region (Turbay et al., 1996), a candidate gene approach was taken to screen the *NKX2-5* gene for mutations in individuals harboring ASD. An identical C to T transition within the homeodomain substituting a methionine for a conserved threonine was found in two families. Separate mutations predicting truncated proteins were found in the remaining two families (Schott et al., 1998). The authors conclude that the homeodomain mutations likely alter either the sequence-specificity or affinity of target DNA binding, suggesting that the CHD in these patients is functionally attributable to haploinsufficiency. It is interesting to note that even within the same kindred, phenotypic variability – ranging from fatal pulmonary atresia with tetralogy of Fallot (TOF) to simple AV block without ASD was seen, demonstrating the extreme phenotypic variability of *NKX2-5* mutations (Schott et al., 1998).

A recent prospective study focused on *NKX2-5* mutations in individuals diagnosed with non-syndromic TOF. Patients were screened for microdeletion of chromosome 22q11 (known to be deleted in many forms of CHD) or other identifiable

cytogenetic abnormality; a positive result during this screen excluded patients from the study. Six *NKX2-5* mutations were identified in the remaining cohort of 114 patients, all of which were found to be outside the homeodomain (Goldmuntz et al., 2001). Asymptomatic first degree relatives of those patients in whom mutations had been identified were found to harbor mutations themselves, indicating that in addition to phenotypic variability, *NKX2-5* mutations are associated with incomplete penetrance in humans.

DiGeorge syndrome (OMIM#188400) and velo-cardio-facial syndrome (VCFS; OMIM #192450) are conditions resulting most often from a common interstitial deletion at human chromosome 22q11. Deletions at this position are a relatively frequent event, detected in 1 of 4000 live births using standard cytogenetic techniques (Baldini, 2002). The hallmark features of DiGeorge syndrome and VCFS include cardiac anomalies (TOF, VSD), characteristic facies and underdevelopment of the thymus and parathyroid glands (Baldini, 2002). Though the phenotypic spectrum is varied, congenital heart defects are seen in approximately 75% of newborns and *del22q11* is perhaps the most common known cause of CHD in neonates (Johnson et al, 2002).

While the underlying molecular events leading to deletion remain undefined, the presence of multiple low-copy repetitive DNA elements has lead to the hypothesis that aberrant recombination between these sequences is the primary mechanism (Stankiewicz et al 2002). The fact that approximately 90% of patients show a stereotypical 3 Mb deletion supports this hypothesis (Lindsay, 2001). It is estimated that this region encompasses 30 genes, and while the phenotype has not been ascribed to a single gene, compelling work in animal models support the notion that haploinsufficiency of the T-

box transcription factor *TBX1* is largely responsible for the phenotype (Lindsay et al., 2001; Jerome et al., 2001; Merscher et al., 2001). The etiology of DiGeorge/VCFS is certainly more complex, given that patients have been identified that carry non-overlapping deletions, and it is suggested that interruption of cis-acting regulatory elements or haploinsufficiency of other genes within the region may be sufficient to produce the phenotype (Baldini, 2002).

Holt-Oram syndrome (OMIM #142900) is a clinically variable autosomal dominant disorder characterized by CHD and radial ray defects. Though there is evidence for genetic heterogeneity (Terrett et al., 1994), most cases of Holt-Oram syndrome are attributable to heterozygous mutations in *TBX5* that predict null-alleles and thus haploinsufficiency of the protein (Bonnet et al., 1994; Basson et al., 1997; Li et al., 1997b). *TBX5* encodes a member of the T-box transcription factor family and its product has recently been shown to associate with NKX2-5 via an interaction between the N-terminal portion of the TBX5 'T-box' and the homeodomain of NKX2-5 (Hiroi et al., 2001). Further, TBX5 and NKX2-5 were found to promote cardiomyocyte differentiation in cell culture, thus placing these two transcription factors within a single cardiac developmental pathway (Hiroi et al., 2001).

Alagille syndrome (AGS, OMIM#118450) is an autosomal dominant condition occurring in approximately 1:70,000 births. Affected individuals have OT abnormalities ranging in severity from peripheral arterial stenosis to TOF. In addition, patients often show developmental defects in the liver, kidney, eye and skeleton (Alagille et al., 1975; Li et al., 1997a; Oda et al., 1997b). Clues to the underlying genetic defect were revealed after *JAG1* was cloned from the Alagille syndrome critical region at 20p12 (Oda et al.,

1997a). Human *JAG1* is the homolog of *Jagged1* in rats, which itself is the homolog of the fly gene *Serrate* (Lindsell et al., 1995). Work in *D. melanogaster* and *C. elegans* had previously shown *Delta* and *Serrate* / LAG2 and APX1 were the ligands for the *Notch*/ LIN12 or GLP1 receptors, respectively (Artavanis-Tsakonas et al., 1995). *Notch* was known to be a highly conserved developmentally essential cell-surface receptor that functions generally to mediate cell fate decisions; among these being the proper specification and formation of mesodermal and neural cell lineages (Corbin et al., 1991; Hoch et al., 1994). *Delta*, *Serrate*, LAG2, APX1 as well as JAG1 are membrane bound extracellular proteins containing a hallmark "DSL" (Delta Serrate Lag-2) cysteine rich region and variable repeats of EGF domains. Recognizing the importance of *Notch* and its ligands in early embryogenesis, Oda and coworkers refined the Alagille critical region and scanned the *JAG1* gene from affected and normal individuals by SSCP. Using this method they were able to find insertions (resulting in frame-shift), deletions and splice donor point mutations. The authors invoke haploinsufficiency, as opposed to a dominant-negative mechanism of pathogenesis given that the phenotype in patients with microdeletions or loss of expression does not differ appreciably from patients harboring point mutations or suffering from alterations in splicing (Oda et al., 1997b). In a related study, heteroduplex migration analysis (HMA) in cDNA heteroduplexes followed by direct sequencing of PCR products was used to screen 10 individuals from four Alagille syndrome families. Small insertions and deletions resulting in a frameshift predicted to cause chain termination were found in each of the four families. It is interesting to note that *JAG1* mutations are analogous to N*KX2-5* mutations in that two affected members of a single family have divergent cardiac phenotypes, again ranging from severe TOF in a

child to a heart murmur reported in her mildly affected father (Li et al., 1997a). Perhaps more strikingly, coding changes in *JAG1* have been identified in non-syndromic TOF characterized by variable expressivity and reduced penetrance, much like a group of *NKX2-5* mutations (Eldadah et al., 2001). This clearly demonstrates that alterations in two separate genes can result in the same restricted cardiac phenotype.

Transcription factors and cell-surface ligands represent but two classes of molecules implicated in the etiology of CHD. Marfan syndrome (OMIM #154700) includes skeletal, ocular and cardiovascular defects as hallmark features. Specifically, mitral valve prolapse, dilated aortic root and a weakened aorta make aortic aneurism leading to dissection one of the leading causes of death in affected individuals (Maslen, 1993). The finding that mutations in fibrillin-1 (*FBN1*) cause Marfan syndrome highlighted the fact that extracellular matrix molecules play a critical role in heart development (Dietz et al., 1991; Maslen et al., 1991).

Williams syndrome (OMIM #194050) is a contiguous gene deletion syndrome marked by variable features including characteristic 'elf-like' facies, short stature, developmental delay and infantile hypercalcemia. However, cardiovascular malformation is one of the most common findings with supravalvular aortic stenosis (SVAS) occurring in the majority of cases. Complete or partial deletion of the 45 kb elastin (*ELN*) gene was found in over 90% of Williams syndrome patients, implicating a second extracellular protein in the occurrence of CHD (Ewart et al., 1993; Nickerson et al., 1995).

While large interstitial deletions as well as mutations within specific genes can lead to congenital heart malformations, so too can the presence of additional genetic

material. This is perhaps best exemplified by Trisomy 21 or Down syndrome (OMIM#190685) which arises from aberrant dosage of genes normally found on human chromosome 21. The stereotypical features of trisomy 21 include congenital heart defects resulting from abnormal endocardial cushion development, mental retardation, characteristic facies, and often hearing loss, though a wider spectrum of clinical features has been defined that includes gastrointestinal anomalies, immune system defects, increased risk of leukemia as well as a reported "Alzheimer's-like" dementia (Korenberg et al., 1994). The occurrence of trisomy 21 increases as a function of advancing maternal age and is reported to range between 1:100 (advanced maternal age) and 1:1000.

Though most often considered in the context of an intact extra copy of chromosome 21, it has been established that specific phenotypic features can be assigned to smaller chromosomal segments. For example, an unbalanced 4;21 translocation in a Japanese pedigree (resulting in del4q35 and duplication of distal chromosome 21q) gives rise to individuals which present with clinical Down syndrome, including typical facial features, heart defects (specifically endocardial cushion defects) and mental retardation (Korenberg et al., 1990). More recently, a gene encoding a neural cell adhesion molecule mapping to human chromosome 21q22.2-22.3 has been isolated (Yamakawa et al. 1998). This gene, called *DSCAM* (Down syndrome cell adhesion molecule) is a member of the immunoglobulin superfamily. Based on developmental expression throughout nervous tissues including the neural tube, cortex and hippocampus and *DSCAM* was initially proposed to contribute to the neurological phenotype associated with trisomy 21. Studies in mouse were able to demonstrate *DSCAM* expression in neural crest derived tissues, including branchial arches, trigeminal ganglia and heart ganglion (Yamakawa et al.,

1998). Using a panel of 19 individuals with partial trisomy 21 and CHD, fine mapping studies were performed in which a Down syndrome CHD candidate region of approximately 5.5 Mb was defined (Barlow et al., 2001). Interestingly, *DSCAM* was found to lie within this narrowed interval and based on its developmental expression in migrating neural crest has been proposed as a candidate gene for CHD associated with trisomy 21 (Barlow et al., 2001). The authors speculate that overexpression of *DSCAM* as a consequence of trisomy may adversely affect EMT and/or cell migration thus leading to abnormal endocardial cushion formation as a possible mechanism.

Congenital heart defects in model systems

While the genetic dissection of human syndromic and non-syndromic CHD has met with some success, the preponderance of current understanding has come from experimental manipulation in both vertebrate and invertebrate systems. One rational approach for modeling human conditions is to make targeted mutations in mouse orthologs of human genes. Using targeted disruption of candidate genes and observing the phenotypic consequences has revealed a broad group of factors, acting alone or in concert during cardiogenesis. Of course, these models are not without their limitations. Nevertheless, considerable progress has been made by the use of this approach and the list of molecules known to be essential for cardiogenesis is increasing at a rapid rate.

Heterozygous *Tbx5* mutant mice were generated to model Holt-Oram syndrome (Bruneau et al., 2001). These mice were found to exhibit subtle changes of the limbs and fully penetrant ASDs (n=7), neither of which were found in wild-type littermates. Furthermore, occasional embryonic death (e16.5) was observed in heterozygous mutants,

which was attributable to intrauterine heart failure. Conduction system disease resulting from a lack of connexin 40 (*cx40*) expression was identified in these animals. Atrial natriuretic factor (*Anf*), normally expressed in atrial myocardium was similarly downregulated. The effect of *Tbx5* mutations on these genes suggested they act in the same pathway. Co- transfection of *Tbx5* and constructs of the putative *cx40* or *Anf* promotors fused to luciferase showed T-box binding elements were necessary to drive reporter expression and were reliant on Tbx5 in a dose-dependent manner. Interestingly, inclusion of an *Nkx2-5* expression construct in these experiments showed an additive effect, demonstrating *Anf* and *cx40* are directly regulated by both transcription factors. In contrast to the heterozygotes, homozygous *Tbx5* mutants expired by e10.5 due to an asymmetrically malformed heart, with striking underdevelopment of left-sided structures and low or absent expression of ventricle-specific markers. Thus, the *Tbx5* knockout mouse has provided additional insight into the pathogenesis of Holt-Oram syndrome and has put *Anf* and *cx40* into a similarly regulated genetic pathway (Bruneau et al., 2001).

The role of *Jag1* was investigated *in vivo* by creating a null allele in the mouse and characterizing both heterozygous and homozygous animals (Xue et al., 1999). As one might predict, interrupting the Notch pathway in homozygous mutants was lethal *in utero*. These embryos were dying at e10.5 of hemorrhage and an apparent lack of large blood vessels in the yolk sac. Similarly, the vascular network of mutant animals was irregular and major vessels had a markedly reduced diameter. In heterozygous mutants, the only apparent defect was a moderately severe ocular phenotype in which the outer edge of the iris was contiguous with the ciliary body musculature. There was no detectable involvement of cardiac, hepatic or skeletal tissues, thus the heterozygotes were

found to be an unsuitable model for Alagille syndrome. The *Jag1*$^{+/-}$ mouse is one example of the sometimes limited and often unpredictable outcome of this technique.

Partial recapitulation of Marfan syndrome was found in mice made deficient for fibrillin-1. These homozygous mutant mice are indistinguishable from wild-type littermates at birth, yet die suddenly between nine and eighteen days of age as a result of cardiovascular complications (Pereira et al., 1997). Interestingly, the original fibrillin-1 null mice did not show the characteristic skeletal defects of Marfan syndrome nor do they show compensatory expression of fibrillin-2. This was likely due to the fact that mutant pups expired prior to weaning from aortic dilatation and dissection (Pereira et al., 1997). A second fibrillin-1 mutant mouse has been generated that resulted from an aberrant targeting event of the original fibrillin-1 targeting vector. This mislaid insertion results in a 5-fold reduction of fibrillin-1 expression without loss of coding sequence and thus does not alter the protein itself (Pereira et al., 1999). Nevertheless, the phenotype of these mice (termed mgR/mgR mice) is relatively benign and homozygous mutants survive into early adulthood. The ability of these animals to survive beyond the neonatal period results in a model of Marfan syndrome which more closely approximates the human condition, including skeletal long bone overgrowth and shortened lifespan resulting from vascular insufficiency (Pereira et al, 1999).

The *Nkx2-5* mutant mouse was produced prior to the discovery of mutations in the human ortholog as causative for non-syndromic ASD and conduction delay. Homozygous mutations were found to result in intrauterine death. These embryos show severe growth retardation and retained linear hearts with no apparent delineation of individual chamber-forming fields. This phenotype is fully penetrant, in that no

homozygous mutants were detected in over 1200 offspring from heterozygous mating. This was a particularly valuable model in that it showed Nkx2-5 is a positive regulator of MLC2V, one of the first proteins specifying ventricular differentiation, thus introducing Nkx2-5 and MLC2V into the same genetic control hierarchy (Lyons et al., 1995).

The *Nkx2-5* mouse is but one example of using a genetic approach to answer questions regarding *in vivo* function. Furthermore, gross phenotypic effects on anatomy or altered viability do not exhaust the utility of such models, as witnessed by the discovery that Nkx2-5 regulates MLC2V production. In fact, deconstructing the genetic pathways essential to vertebrate heart assembly has been significantly advanced by perturbing genes suspected to play a role in cardiogenesis one by one. Perhaps nowhere else is this more evident than the relationships that have been established among transcription factors.

The role of *Mef2c* in cardiogenesis was investigated in a mouse model (Lin et al., 1997). *Mef2c* is normally expressed throughout the heart tube uniformly. While the heterozygous mice were indistinguishable from wild-type animals, homozygous mutants died between e9.0 and 10.5. *Mef2c* homozygous mutants produced a linear heart tube that began contractions, however the contractions were weak and irregular. Further, the heart tube did not undergo rightward looping, a morphological shift necessary to bring the presumptive atria and ventricles into alignment. Other cardiogenic genes were investigated in these mutants. In direct contrast to *Nkx2-5* mutations, *Mef2c* mutations were found to have no effect on Mlc2v production or pattern of expression. One simple explanation for this observation is that *Nkx2-5* and *Mef2c* regulate different genetic pathways in heart. Furthermore, the consequences of *Mef2c* mutations on the basic Helix-

Loop-Helix (bHLH) transcription factors dHAND and eHAND were probed. These proteins are known to be expressed in a complimentary fashion throughout heart forming tissue, with dHAND localized predominantly to the right side of the heart, and eHAND primarily expressed in the conotruncus and future left ventricle (Srivastava et al., 1997). Transcripts of *dHAND* were found to be downregulated at the time of looping and the right ventricle failed to form, whereas *eHAND* expression was expanded to encompass the entire heart tube. *Nkx2-5* expression appeared similar to wild-type, again suggesting *Nkx2-5* is functioning in a separate pathway. An alternative explanation is that *Nkx2-5* acts upstream of *Mef2c*.

When *dHAND* mutant mice were generated, again the heterozygotes did not show any obvious defects. No live-born homozygous mutants were obtained from 155 offspring indicating embryonic lethality. Inspection of embryos recovered by Caesarian sections revealed death between embryonic days 10 and 11. In these mice, cardiac looping does begin, but arrests abruptly as a single left ventricle forms a connection with the outflow tract. No evidence for a right ventricle was detected (Srivastava et al., 1997). Also found was an absence of aortic arch arteries and a dilated aortic sac. These structures are populated with migrating cardiac neural crest during development. The abnormal anatomy in these structures indicates *dHAND* expression, in addition to a neural-crest component is required for proper formation. *In situ* hybridization experiments using an *eHAND* probe detected robust signal throughout the heart, which further suggests *dHAND* plays a regulatory role in *eHAND* expression. Finally, while the expression patterns of many of the genes discussed previously (*Anf, Mlc2v, Nkx2-5*) appeared normal, the expression of the cardiogenic transcription factor Gata4, was

largely eliminated in *dHAND* mutant hearts, which places *Gata4* downstream of *dHAND* and in the same genetic module, either directly or indirectly (Srivastava et al., 1997).

The study of homozygous *eHAND* mouse mutants also revealed embryonic lethality, only developmental arrest was found between embryonic days 8.5-9.5 (Firulli et al., 1998; Riley et al., 1998). Primary defects were found in the yolk sac and heart, the latter of which never underwent looping. In one mouse model, tracking *eHAND* expression was facilitated by the use of a 'knock-in' β-gal reporter construct that replaced the native start site with that of the reporter (Firulli et al., 1998). *eHAND* expression was localized to the left ventricle and outflow tract, as well as extraembryonic mesoderm. Additionally, reporter signal was detected in the first branchial arch. Again, expression of *Anf*, *Mlc2v*, *Mef2c* and *Nkx2-5* were unaffected in these models, suggesting *eHAND* lies downstream of these elements in development. Presumably, the early embryonic death results not from cardiac defects directly, but a lack of nutrients due to malformed extraembryonic components, including chorion, amnion, allantois and yolk sac (Firulli et al., 1998). This was confirmed by tetraploid-rescue of chimeric embryos, which survived until at least until e9.5, at which point rightward looping failed (Riley et al., 1998). In these rescued embryos, *Gata-4*, *Nkx2-5*, *Mef2c* and *dHAND* were expressed normally. However, *Mlc2v* transcripts were not detected, suggesting *eHAND* is necessary for maintenance of expression of this gene. Additional mutant mice exhibiting a cardiovascular phenotype are listed in Table A.

Table A: Knockout mice with cardiovascular defects

GENE	CLASS	HEART PHENOTYPE	REFERENCE
Gata4	Transcription factor	Aberrant heart field migration; failure to form heart tube; embryonic lethal	(Kuo et al., 1997; Molkentin et al., 1997)
Madh6 (Smad6)	Signal-transduction/ transcriptional co-activator	OT septation defects; valve hyperplasia; survive to adulthood	(Galvin et al., 2000)
Neuregulin	Soluble ligand (paracrine factor)	Abnormal endocardial cushions and trabeculae; embryonic lethal	(Gassmann et al., 1995; Meyer and Birchmeier, 1995)
Connexin43	Intercellular membrane channel/ gap junction protein	OT enlargement with severe blockage; post-natal lethality	(Reaume et al., 1995)
N-cadherin	Cell adhesion	Expanded pericardial cavity, failure to develop normal myocardium; embryonic lethal	(Radice et al., 1997)
Versican/*hdf* (random insertion of transgene lead to heart defect mouse)	Extracellular matrix glycoprotein	Absent endocardial cushions; primitive right ventricle and OT fail to form; embryonic lethal	(Mjaatvedt et al., 1998)
m-Bop	Histone-deacetylase dependent transcriptional repressor	Extreme thickening of cardiac jelly (ECM); failure of right ventricle formation embryonic lethal	(Gottlieb et al., 2002)

3p- syndrome

Congenital abnormalities of the heart frequently result from chromosomal aberration resulting in an altered dosage of one or more genes (Maslen, 1996). 3p- syndrome is characterized cytogenetically by monosomic deletion of the distal portion of 3p25-pter, or by interstitial deletion at 3p25-26. Since the initial report in 1978, fewer than 25 cases have been presented, nearly all of them sporadic, demonstrating the rarity of this disorder (Verjall, 1978; Phipps et al., 1994). The hallmark features of 3p- syndrome include low birth weight, pre- and postnatal growth delay, psychomotor and mental retardation, hypotonia, microcephaly, ptosis, low set malformed ears, micrognathia, telecanthus and long philtrum. Differences in proximal chromosomal break points result in phenotypic variability including postaxial polydactyly, cleft palate, renal and gastrointestinal anomalies, rockerbottom feet, seizures, triangular face, preauricular pits and impaired hearing. Congenital heart defects are also a variable feature associated with 3p- syndrome. Approximately one-quarter of all infants with 3p- syndrome expire, with CHD as the principle cause (Phipps et al., 1994; Drumheller et al., 1996).

The underlying molecular genetics and genotype-phenotype correlation of 3p- syndrome have been examined. One individual with an interstitial deletion of 3p25-26 and the characteristic features of the syndrome was studied. The chromosomal breakpoints in this patient were between D3S1317 and D3S17, suggesting that the minimal critical region must lie within this 21cM interval (Mowery, 1993). Using a combination of fluorescence *in situ* hybridization (FISH) and analysis of polymorphic DNA markers from chromosome 3p25-26, the chromosomal breakpoints and phenotypic

consequences were described in five additional patient cell lines. It was found that patients with the most proximal deletions presented with cardiac septal defects, whereas patients with more distal deletions lacked CHD, suggesting haploinsufficiency of one or more genes in this region leads to heart malformation. Further, through these analyses the authors determined that a gene for normal cardiac development must lie between markers D3S1250 and D3S18 (~4 cM) (Phipps et al., 1994). In a subsequent study, using somatic hybrid cell lines generated from 3p- syndrome patients, derivative 3 chromosomes were isolated. PCR analysis using microsatellite markers from the region implicated in CHD refined the critical region. An approximately 1 Mb interval bounded by markers D3S1585 and D3S1263 was reported. Interestingly, the most severely affected patient in this study group had the most extensive deletion and was reported to have an endocardial cushion defect (Drumheller et al., 1996).

CRELD1

We have discovered and characterized a gene mapping to human chromosome 3p25 (Rupp, Fouad et al., 2002; Appendix III). The sequence characteristics of this gene reveal similarity to fibrillin1 and 2 as well as other matrix proteins. Our group obtained a 2.1 kb cDNA clone from the Whitehead Institute (WI11041). Complete sequencing revealed a 1263 bp open reading frame encoding a protein of 420 amino acids. Originally this protein was named 'cirrin' (Maslen, 1999), however the Human Genome Nomenclature committee has recently designated this protein CRELD1 (cysteine-rich with EGF-like domains 1), and thus cirrin will be referred to as CRELD1 hereafter.

A bacterial artificial chromosome (BAC 172I17) has been identified that contains the human genomic sequence. Intron-exon organization has been determined by direct sequencing of PCR amplified fragments derived from BAC172I17. This was confirmed by direct sequence analysis of PCR products derived from human genomic DNA. The *CRELD1* gene has 11 coding exons and spans approximately 12 kb. There is direct evidence for variable splicing of *CRELD1*, with one of two alternative terminal exons utilized.

Analysis of the amino acid sequence of CRELD1 using publicly available online tools has revealed motifs within the protein with similarity to elements found in extracellular molecules. A secretion signal is predicted at the amino-terminus followed by a short, poorly conserved proline-rich region. Adjacent to the proline-rich region, a sequence enriched for tryptophan and glutamic acid has been ascertained (the WE domain). This sequence, spanning 135 amino acids, seems to be unique to *CRELD* family members. A pair of EGF domains, followed by a pair of calcium binding EGF-like domains are positioned after the WE domain. Two membrane-spanning helices are predicted by computer analysis, indicating CRELD1 likely extends into the extracellular space yet remains anchored to the cell membrane. This model further predicts the carboxyl-terminus of the protein lies outside the cell, as shown in figure 2.

Figure 2

Schematic representation showing CRELD1 domain structure. The WE domain (yellow) extends into the extracellular space. A pair of EGF domains and calcium binding EGF domains (blue) lie adjacent to the membrane. Predicted transmembrane spanning helices are shown in red; carboxyl terminus is extracellular.

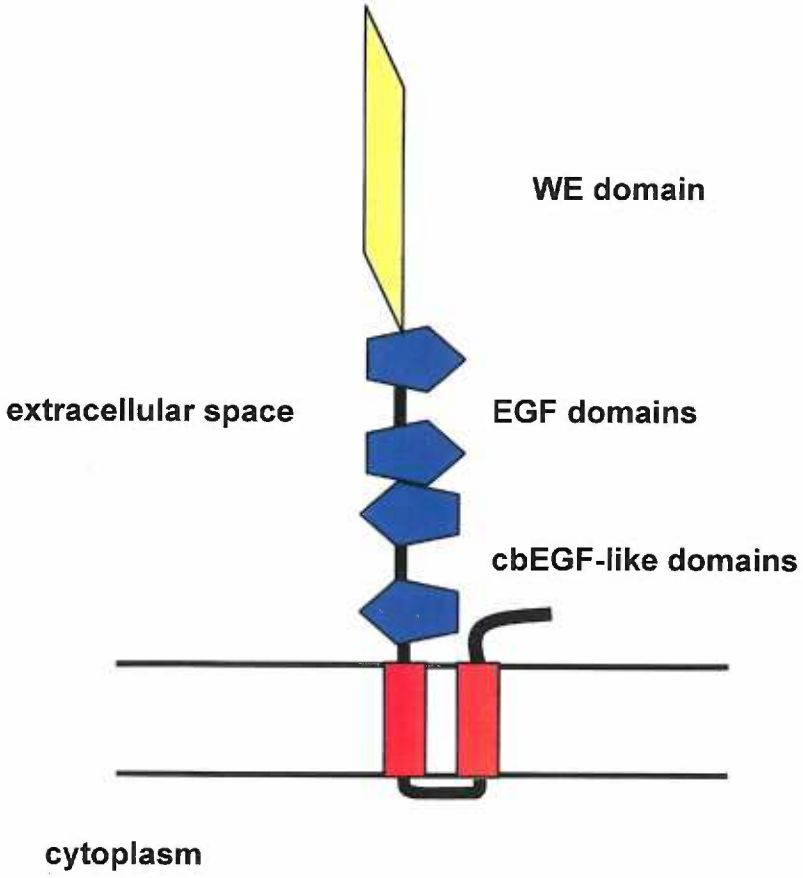

WE domain
extracellular space
EGF domains
cbEGF-like domains
cytoplasm

Figure 2

Multiple-species Southern blot analysis under stringent conditions has revealed that *CRELD1* is found in mammals and birds, though not in yeast. Commercially available adult and fetal multiple-tissue Northern blots have been probed with a *CRELD1* cDNA fragment. A 2.1 kb transcript was found to be widely expressed, with particularly high levels found in fetal and adult heart as well as skeletal muscle. Whole mount *in situ* hybridization studies in chick demonstrate expression in the endocardial cushions and myocardium of the developing heart, as well as in the limb buds, developing brain, mandible and branchial arches (Rupp, Fouad et al., 2002).

CRELD1 encodes the founding member of a new family of proteins belonging to the EGF superfamily, and a second closely related gene (*CRELD2*) has been identified. *CRELD2* ESTs have been localized to chromosome 22p13 (GenBank accession number NP_077300) (Rupp, Fouad et al., 2002). While CRELD2 shares significant sequence (38% identical, 51% similar at the amino acid level) and structural domain similarity with CRELD1, it is predicted to be secreted into the extracellular space and contains no putative membrane spanning domains.

Cell lines from 3p- syndrome patients have been obtained from the NIGMS Human Mutant Cell Repository (Coriell Institute). Two of these cell lines (GM10922 and GM07873) are derived from patients with endocardial cushion defects. Cell line GM10985 is from a patient without heart defects. The patients represented by these cell lines were carefully evaluated clinically for the standard features of 3p- syndrome prior to inclusion in the repository. Fluorescent *in situ* hybridization (FISH) studies have been performed on metaphase chromosome from these cell lines using a 11kb *CRELD1* probe and a chromosome 3 alpha satellite probe to confirm chromosome identity. FISH results

for cell lines GM10922 and GM07873 consistently show two chromosome 3 alpha satellite signals, but only one *CRELD1* signal, demonstrating *CRELD1* is deleted in those patients with CHD. FISH results for cell line 10985 were ambiguous in that a positive signal for *CRELD1*, while present, was not consistently detectable (Rupp, 1999).

Hypothesis

We propose that CRELD1, a newly characterized and highly conserved extracellular protein that is expressed at high levels in the endocardial cushions and myocardium of the developing heart, is critical for normal cardiac morphogenesis. More specifically, we propose that CRELD1 plays a key role in the formation of valve leaflets and the ultimate structural alignment of the internal septa based on its expression pattern. Defects in this process are predicted to result in AVSD. We have undertaken a project designed to question CRELD1 involvement in heart development by seeking to recapitulate the endocardial cushion defects (AVSD) seen in CHD-positive 3p- syndrome patients by generating a targeted *Creld1* haploinsufficient mouse. As an ancillary approach, we further intend to probe the function of CRELD1 by conducting a yeast two-hybrid library screen in an attempt to identify proteins with which CRELD1 interacts.

Chapter II

Materials and Methods

Materials and Methods

Mouse genomic library

In order to isolate a DNA fragment of murine *Creld1* to use as a reagent for constructing a gene-targeting replacement cassette, a mouse strain 129 genomic library was obtained from the OHSU Transgenic Mouse Core Facility. The titer of the library was reported to be 1.0×10^7 pfu/mL and was packaged in Stratagene Lambda FixII with an average insert size ranging between 15 and 20 kb. The library was estimated to contain 1.1×10^6 independent phage.

Genomic library titering

A streak of *E. coli* host strain XL-1 Blue MRA was made on an LB plate following standard techniques. Individual colonies were then re-streaked onto additional LB plates followed by 37 °C incubation for 12 hours. A single colony from the working stock plate was used to inoculate 50 ml LB broth supplemented with 0.2% maltose and 10 mM $MgSO_4$. The culture was incubated over night at 30 C while shaking. The cells were then collected aseptically (10 min at 2000 RPM) and the pellet resuspended in 15 mL 10mM $MgSO_4$. The cells were diluted to $OD_{600} = 0.5$ in 10mM $MgSO_4$ and stored at 4 °C for up to 24 hours. Serial dilutions (1:10; 1:100; 1:1000) were made of the library-containing bacteriophage in 1X SM buffer (10X Stock: 5.8 g NaCl, 2.0 g $MgSO_4$ /$7H_20$, 50 ml 1 M Tris-HCl (pH 7.5), 5.0 ml 2% gelatin; qs to 1 liter with d H_20 and sterilize). One µL representing each dilution of library was added to 200 µL host cells

and incubated 15 min at 37 °C. Following incubation, the mixture was added to 2.75 mL NZY top agar and plated on 100mm NZY plates then incubated at 37 °C overnight.

Plaque transfer to nylon membrane

A fresh culture of *E. coli* host strain XL-1 Blue MRA was prepared as described for genomic library titering. To screen a minimum of 1 X 10^6 plaques, 20 plates containing 50,000 pfu/plate were prepared. 34 µL bacteriophage (2.9 X 10^7 pfu/mL) were incubated with 12 mL host cells for 15 min at 37 °C with gentle agitation. 600 µL cell suspension was then mixed with 6.5 mL prewarmed NZY top agar and poured carefully over 150 mm LB agar plates, followed by incubation at 37 °C for 16 hours. Positively charged nylon membrane disks were labeled in pencil. Duplicate lifts were performed sequentially per plate by placing the membrane directly onto the agar for 2 to 3 min per disk. Orientation marks were made and the disk was removed from the surface of the plate then transferred to a Whatman 3MM filter saturated with 0.5M NaOH, plaque-side up for 2 to 3 min. Disks were then washed in excess 5X SSC for 1 min and air-dried. Library DNA was immobilized on the disks by baking at 80 °C X 2 h under vacuum and stored at room temperature.

Preparation of genomic probe

Human *CRELD1* cDNA was used to search the mouse EST database. Seventeen unique overlapping murine EST's were found that shared significant sequence homology with human cDNA. Primers were designed in areas of greatest similarity. These primers were added to a PCR reaction consisting of 100 ng mouse strain 129 genomic DNA as

template and Gibco BRL *Taq* DNA Polymerase (Rockville, MD, cat. #18038) reagents. The reaction conditions were as follows (note: all concentrations are initial):

68 µL dH_2O
10 µL 10X Reaction buffer
8 µL dNTPs [10 mM]
4 µL Forward Primer E1-2F [25 µM]
4 µL Reverse Primer E1-2R [25 µM]
4 µL $MgCl_2$ [50 mM]
1 µL Taq DNA Polymerase
1 µL Mouse genomic DNA
100 µL total

Conditions: 95 °C X 3'
[94 °C X 45" / 61 °C X 45" / 72 °C X 2'] X 35 cycles
72 °C X 5'

The PCR product derived from this reaction was electrophoresed on a 1% agarose gel prepared in 1X TBE containing 1 µg/mL ethidium bromide and visualized via UV transillumination. A band migrating between 400 and 600 base pairs (bp) was purified from the agarose using QIAquick Gel Extraction Kit (Qiagen Inc., Valencia CA Cat.#28706) following kit protocol. The concentration of the extracted product was estimated by loading 5 µL on a 0.8% agarose gel and comparing with HyperLadder quantitative standard (ISC BioExpress, Kaysville, UT Cat# C-5087-200). Approximately 20 ng of the recovered fragment was cloned using pMOS*Blue* blunt ended cloning reagents according to kit protocol (Amersham Pharmacia Biotech, Buckinghamshire UK Cat.# RPN5110). Ligation products were transformed into MOS*Blue* competent cells following standard methods and plated on LB Agar plates containing 50 µg/mL ampicillin and 15 µg/mL tetracycline then incubated at 37 °C X 16 hrs. Colonies were used to inoculate 3 mL LB containing 50 µg/mL ampicillin, which were incubated at 37

°C overnight shaking at 220 RPM. pMOS plasmids were recovered from MOS*Blue* cultures using QIAprep Spin Miniprep columns (Qiagen, Cat #27106) following manufacturer's instructions. Isolated plasmids were then screened using primer set E1-2F/R as described above. Plasmids yielding PCR products of expected size were sequenced at the Shriners Hospital Sequencing Core facility using vector primer T7 (Appendix I).

Labeling for chemiluminescent detection

Gel purified PCR product (50 ng) was labeled using ECL Random Prime Labeling and Detection System (version II) (Amersham Pharmacia Biotech, Buckinghamshire UK Cat.#RPN3040/3041) following system procedures. Dilutions of labeled probe were made (1:5, 1:25, 1:50, 1:100, 1:250, 1:500, 1:1000 in TE) and determination of probe labeling efficiency was made using a rapid labeling assay. This was accomplished by applying 5µl aliquots to positively charged nylon filters followed by a gentle wash in 2X SSC (60°C) X 15 min and visualization by UV transillumination and comparison against standards.

Probe hybridization, stringency washes and signal detection

Filters were divided into groups of 10 and hybridized with labeled probe. Hybridization buffer was prepared according to kit instructions. Labeled probe was denatured at 100°C for five min and added to buffer. Hybridization was performed overnight at 60°C. Two stringency washes were carried out: 1XSSC/0.1% SDS followed

by 0.5X SSC/0.1% SDS each at 60°C for 15 min. Filters were then rinsed briefly in excess buffer A (100 mM Tris-HCl, 600 mM NaCl pH 7.5) and incubated an additional 30 min at room temperature using supplied blocking reagent. A 1:1000 dilution of α-fluorescein-HRP conjugate was prepared in buffer A containing 0.5% BSA (Sigma, St. Louis, MO Cat.# A-2153) in a total volume approximately equivalent to 0.25ml/cm^2 of membrane. Filters were incubated at room temperature for 30 min followed by 3 X 10 min washes in excess buffer A containing 0.1% Tween-20. Kit supplied detection solution (0.125ml/ cm^2 membrane) was prepared according to protocol and applied directly to filters. After 1 minute, excess reagent was removed and filters were wrapped in saran-wrap. Filters were exposed to Amersham Hyperfilm ECL (Cat.#RPN3114) for 10 min at room temperature and developed.

Identification and isolation of plaques

Films showing pinpoint signals were aligned with original plates and signals present in duplicate were identified. Plates were placed over film and the stack was viewed over a light box. Using sterile pasture pipettes, plugs of agar containing viral plaques were isolated. Agar plugs were placed in 1 mL 1X SM buffer and bacteriophage were eluted overnight at 4°C. Eluant was diluted 1:50 in 1X SM buffer and used to inoculate additional XL-1 Blue MRA *E. coli* . These cultures were used to prepare additional plates that were then re-probed as described above to identify single, well-isolated plaques.

Purification and subcloning of murine genomic DNA

A 5 mL culture of XL-1 Blue MRA *E. coli* was incubated overnight in LB containing 0.2% maltose; 10 mM $MgSO_4$. 500 µL of the overnight culture was inoculated with 25 µL viral eluate and incubated for 20 min at 37 °C. This was then added to 100 mL prewarmed supplemented LB in a sterile flask and incubated at 220 RPM at 37°C. The culture was monitored until lysis (defined as rapid clearing of the otherwise turbid culture) occurred. Viral DNA was isolated from this solution using the Wizard Lambda Prep DNA Purification System according to manufacturers instructions (Promega, Madison, WI Cat#A7290). Purified DNA was assessed by PCR (primers E1-2F/R) as described previously. PCR positive clones were digested with *NotI* following supplier's recommended conditions in order to excise the murine insert. *NotI* was similarly used to linearize pBluescript II SK (-) (Stratagene, La Jolla, CA Cat#212206) for use in subcloning and subsequent restriction mapping of the genomic fragment. *NotI* restricted bacteriophage insert and pBluescript II SK(-) were gel purified as described previously. The vector was treated with calf-intestine alkaline phosphatase according to supplier's suggested conditions (Boehringer Mannheim, Indianapolis, IN Cat.#713 023). Ligation reactions were performed using a 1:10 ratio of vector: insert in a 20 µL reaction using T4- ligase and incubation overnight at 15°C . Aliquots (1 µL and 5µL) of the ligation reaction were then used to transform competent DH5α *E. coli* using standard methods (Ausubel, 2001). Colonies were picked and grown under selection followed by plasmid recovery using the alkaline lysis miniprep method (Ausubel, 2001). Individual plasmids were assessed for insert using *NotI* restriction digest followed by electrophoresis and visualization as described previously. Sequencing reactions were

prepared using primers E1-2F, E1-2R, E2-3F and E2-3R (Appendix I). Sequencing was performed on an ABI 373 Sequencer at the Vollum Sequencing Core Laboratory. Successful sequencing reactions were used to design additional primers for further clone characterization. Initially, these included MExon1.Seq and MIntron1.seq (Appendix I). Upon positive identification of clones, maxi preps (Qiagen) were performed to obtain ample DNA for characterization by restriction mapping. Resulting plasmid was extracted using standard phenol-chloroform methodology. Following extraction, sample was precipitated with sodium acetate and ethanol, followed by resuspension in 10mM Tris pH 8.0.

Radiation hybrid mapping of murine *Creld1*

A mouse/hamster whole genome radiation hybrid panel was obtained through Research Genetics (McCarthy et al., 1997). The panel contains DNA extracted from 100 mouse/hamster clones, with mouse genome representation in the hybrid array estimated at between 20-25%. Using primer pair E1-2F/R and Advantage 2 PCR reagents, a single 528 bp fragment was obtained from 100 ng mouse strain 129Sv genomic DNA but not 100ng hamster genomic DNA (control). Two-step cycling parameters were used: 95 °C for 1 min. initial denaturation followed by 95 °C for 30 s then 68 °C for 1 min for 30 cycles. Using these conditions, 50 ng DNA from each hybrid was used in a 25 µL PCR reaction, followed by visualization of 15 µL reaction product on a 0.8% agarose gel containing 0.05 µg/mL ethidium bromide in 1X TAE running buffer. The panel was screened and scored in duplicate. Total mouse and hamster genomic DNA controls were included. A single ambiguous result was scored "R", whereas all positive results were

scored "1" and negative "0". The data set was analyzed using the online Jackson Laboratory T31 Mouse Radiation Hybrid Database found at: (http://www.jax.org/resources/documents/cmdata/rhmap/rhsubmit/html).

Localization of restriction sites on the murine *Creld1* clone

In order to generate a partial restriction map of the genomic clone, the insert was excised from pBluescript II SK(-) using *NotI* followed by precipitation using 3M sodium acetate and 70% ethanol (Ausubel, 2001). A panel of enzymes was selected for restriction of excised insert [*SacI, SacII, SmaI, BamHI, HindIII, EcoRI, EcoRV, XbaI, PstI, SpeI, ClaI, SalI, XhoI, BstXI*; various suppliers]. 1 µg of DNA was used in each reaction following recommended conditions for each individual enzyme. Products were electrophoresed and visualized as described previously, followed by transfer to positively charged nylon membrane with an alkaline buffer (Ausubel, 2001). Oligonucleotide probes were end-labeled with T4 polynucleotide kinase (Roche Molecular Biochemicals, Mannheim, Germany Cat#174 645) in a 30 µL reaction (all concentrations initial):

3	µL	oligonucleotide (50 pmol)
3	µL	10X Polynucleotide Kinase buffer
5	µL	[γ^{32}P] dATP (10 µCi/µL)
0.5	µL	T4 Polynucleotide Kinase (10U/µL)
18.5	µL	dH_2O

Incubate at 37 °C X 30 min; Heat inactivate at 65°C X 5 min.

Unincorporated $\gamma^{32}P$ dATP, salts and enzymes were removed from the probe using Microcon YM3 spin columns (Millipore Corporation, Bedford MA Cat#42403). This was accomplished by adding 200 µL TE to the sample and centrifuging the column at 13,000 RPM X 100 min followed by washing the column with an additional 200 µL TE at 13,000 RPM X 45 min. Probe was collected by inverting the column and pulsing at 1000 g X 3 min. Specific activity of each probe was determined by adding a 1:100 dilution of 1 µL purified probe to 10 mL scintillation fluid followed by detection. SDS pre/hybridization buffer for use with oligonucleotides was prepared (2X solution):

25 mL 2M $NaHPO_4$ pH7.2

0.2 mL 0.5 M EDTA

4 mL 25% (w/v) BSA

20 mL 25% SDS

Working solution is prepared by adding 12 mL 40% formamide and 3 mL dH_2O to 15 mL 2X SDS pre/hybridization buffer. Blots were prehybridized for a minimum of 15 minutes at 42 °C. Denatured probe (5 min @ 95°C) was added to hybridization solution and allowed to hybridize over night at 42°C with mild agitation. Following incubation, blots were washed in 2X SSC/0.1% SDS for 15 min at room temperature followed by a 1 min rinse in 0.2X SSC/0.1% SDS at 25°C. The membranes were then wrapped in saran-wrap and exposed to x-ray film for 24-72 hours. Blots were stripped for reuse with successive probes by placing membranes into a boiling solution of 0.1% SDS and shaking gently until the solution had cooled to room temperature.

Generation of the knockout-replacement construct

A strategy for assembling a targeting vector suitable for homologous recombination was designed. In order to accomplish assembly, a 10.7 kb genomic fragment of the mouse *Creld1* gene was isolated by restricting the full-length genomic clone with *SalI* and *SpeI* according to standard methods. Following gel purification, this fragment was subcloned into the pEGFP (Clontech, Cat# 6077-1) vector backbone using routine ligation conditions. Upon isolation of a suitable subclone, a 1.9 kb fragment was excised from within the genomic insert using *BamHI* and *SacI* following manufacturer's recommended conditions; the resulting plasmid was gel purified using QIAquick Gel Extraction Kit. A pair of oligonucleotides containing a *BglII* site (Linker "T" and "B"; Appendix I) was then designed for insertion into the subcloned 10.7 kb fragment between existing *BamHI* and *SacI* sites. Double-stranded linker was prepared by combining 0.5 µg each oligonucleotide in 50mM Tris Buffer (pH 8.0) in 10 µl total volume. The solution was heated at 65 °C for 5 min and allowed to cool and remain at room temperature for 1 hour to allow annealing. Prepared linker solution was diluted 1:500 in dH_2O and used in standard ligation reactions, which were performed by combining 1, 2 and 4 µl linker respectively in 10 µl total volume and incubating at 15 °C overnight. Ligation products were transformed into competent DH5α *E. coli* and cultured on appropriate media. Resulting colonies were grown in 3 mL cultures for 12-16 hours and plasmids were recovered as described previously. Clones containing appropriate linker placement were identified by restriction with *BglII* followed by electrophoresis and visualization according to protocol. The vector pACN-1 was obtained as a gift from Dr. Kirk Thomas, University of Utah (Bunting et al., 1999). pACN-1 was used as starting

material from which a "self-excising" neomycin resistance cassette was isolated for use in the targeting construct as a positive selectable marker. This was accomplished by treating pACN-1 with *BglII* under standard conditions followed by gel purification of the desired fragment as described previously. The purified *neo*-cassette was inserted into the linker positioned within the mouse *Creld1* clone as detailed above (figure 5, results). Orientation of the *neo*-cassette was deduced by using *HindIII* to fragment the construct and confirming one of two possible outcomes made *a priori.* Next, a negative selectable marker was also incorporated downstream of the 3' arm of homology. The vector pMC1-TK was obtained as a gift from Dr. Scott Stadler of Shriners Hospital for Children in Portland, OR, from which a thymidine kinase gene was isolated using *XhoI* and *SalI* according to protocol. *XhoI* and *SalI* have complementary cohesive ends, which allowed for easy incorporation of TK into a pre-existing *SalI* site that was conveniently located adjacent to the 3' end of the mouse genomic sequence. Upon complete assembly of the replacement construct and confirmation of structure, a maxi-prep was performed (Qiagen). 60 µg of plasmid was linearized with *SalI* (recreated now downstream of the TK gene) according to standard conditions. The linear construct was precipitated using 3M sodium acetate and 70% ethanol as described (Ausubel, 2001) and resuspended in sterile TE under aseptic conditions. The complete, linear replacement vector was then delivered to the OHSU Transgenic / Gene Targeting Facility for electroporation, culture and selection of ES cells. Later, 200 µg of targeting vector was prepared in a similar fashion for submission to the Gene Targeted Mouse Service at the University of Cincinnati, Dept. of Molecular Genetics, Biochemistry and Molecular Biology (231 Albert Sabin Way, Cincinnati, OH 45267).

Partial analysis of BAC inserts

In order to develop a Southern blot assay suitable for identifying positively targeted ES cells, BACs were obtained from Incyte Genomics (St. Louis, MO) using Incyte BAC Mouse II PCR library screening service. Via this service, three individual BACs were identified that yielded appropriately sized PCR product using primers E1-2F/R as detailed above. Agar stabs of BACs (pBeloBAC11 vector) were used to make streaks on LB agar plates containing 12.5 µg/mL chloramphenicol. BACs were isolated from 500 mL culture using components and methods as outlined in the Qiagen Large Construct Kit Handbook (Cat#12462). Upon isolation of BAC DNA, inserts were characterized by PCR. In addition to primer pair E1-2F/R, several additional primers were used. These included E2-3F.2 and E2-3R, M.Ex3-4F and M.Ex3-4R as well as M.Ex5-6 F and R (Appendix I). A two step PCR protocol was followed using Advantage 2 PCR components (1X reaction, all concentrations initial; Clontech Cat#K1910-1):

	38 µL dH_2O
	5 µL 10X Reaction buffer
	1 µL 50X dNTP mix [10 µM each]
	2 µL Forward Primer [25 µM]
	2 µL Reverse Primer [25 µM]
	1 µL 50X Advantage 2 Polymerase Mix
	1 µL BAC template (50 ng/µL)
	50 µL total
Conditions:	95 °C X 3'
	[95 °C X 30" / 68 °C X 1 min] X 30 cycles
	68 °C X 1 min

PCR products were obtained, gel purified and directly sequenced on an ABI 373 DNA Sequencer. Because primer set E2-3F.2/E2-3R spanned the 3' end of the genomic clone

used to make the targeting vector, additional primers were designed from this relatively large PCR product for purposes of obtaining additional sequence and deriving probes suitable for Southern analysis. These include primers E2-3F.3, E2-3F.4 and E2-3F.5 (Appendix I). Similarly, using sequence from the same PCR fragment, primers Neo1R and Neo1R2 (Appendix I) were designed anticipating a PCR genotyping assay for rapid identification of the null allele in mice. In order to generate PCR products from BAC DNA for assembly into a complete contig spanning the 3' end of the genomic clone through mouse exon 5, several permutations of the primer pairs listed above were utilized, including M.Ex3-4F and M.Ex5-6R. Primers E2-3F.4 and E2-3F.5 were also paired with M.Ex3-4R in order to generate smaller fragments in regions presupposed suitable for probes. Furthermore, a partial restriction map of the contig was constructed using a combination of direct sequence analysis using the online sequence analysis tool Webcutter 2.0 (http://www.firstmarket.com/cutter/cut2.html) and enzyme treatment as detailed previously.

Characterization of murine *Creld1* cDNA

Using sequence generated from PCR products derived from BACs and mouse genomic DNA, an EST likely to contain full-length murine *Creld1* was identified from the EST database (blastn parameters with the "est_mouse" database specified; http://www.ncbi.nlm.nih.gov/BLAST/). An EST representing and IMAGE Consortium Clone ID # 1547690 (GenBank accession #BE200505) was obtained from Incyte Genomics Inc. The cDNA had been cloned into vector pT7T3D-Pac and the insert ends

were thus sequenced using standard T7 and T3 sequencing primers and conditions. Additionally, sequence generated using primers M.Ex3-4F, M.Ex5-6F and M.Ex8-9F (Appendix I) was used to complete assembly of the insert sequence.

Extraction of ES cell genomic DNA

Samples from the OHSU facility were received as ES cell colonies in duplicate 96 well plates. Due to a limited number of cells in each well, all DNA extraction steps as well as treatment with restriction enzyme were performed directly in the plates. DNA was extracted according to a method developed in the laboratory of Allen Bradley (Ramirez-Solis et al., 1992). ES cell pellets were received in 1.5 mL microfuge tubes from the University of Cincinnati. Pellets were suspended in 0.5 mL 1X lysis buffer (100 mM Tris-HCl pH 8.3, 5 mM EDTA, 200 mM NaCl; 0.2% SDS [final concentration] and 100 µg/mL Proteinase K added immediately before use) and incubated overnight at 37 °C. DNA was precipitated by adding 1.0 mL ethanol. Tubes were centrifuged 1 min and the supernatant was carefully aspirated. Pellets were washed in 1 mL cold 70% ethanol and again the supernatant was removed. Remaining liquid was drawn up with a pipetman. Pellets were air dried for a minimum of 10 minutes and resuspended by adding 150 µl 10 mM Tris pH 8.5 and incubating overnight at 50 °C.

Southern analysis of mouse genomic DNA

A strategy to screen ES cells (or high molecular weight genomic DNA from tail clips and thus assess germline transmission) for appropriate insertion of the targeting vector was devised. A 757 bp probe was generated by PCR (primer pair E2-3F5/M.Ex3-4R) using BAC DNA as template with Advantage 2 reagents and conditions listed above. 50 ng gel purified probe was labeled using [α^{32}P]dCTP (NEN Life Sciences, Boston Cat#BLU513H) and "Ready-To-Go" DNA Labeling Beads (Amersham Pharmacia Biotech, Cat#27-9240-01) following manufacturer's instructions. Probe was purified using QIAquick Nucleotide Removal Kit components and methods (Qiagen Cat#28304). Specific activity of each probe was determined by adding a 1:100 dilution of 1 µL purified probe to 10 mL scintillation fluid followed by detection. Approximately 5 to 8 µg genomic DNA template was treated with *BglII* and allowed to incubate at 37°C for 12 h in a 50 µl reaction after which 5 µl 10X loading dye was added. A 0.8% agarose gel containing 0.05 µg/mL ethidium bromide was prepared in 1X TAE. For each blot, approximately 40 µl sample was loaded and electrophoresed at 60V X 3 h. The gel was removed, rinsed briefly in dH_2O and immersed in 0.25M HCl for 25 min. Next, the gel was placed in 0.4M NaOH for 30 min then DNA was transferred to a positively charged nylon membrane over night using 0.4M NaOH as the transfer buffer (Ausubel, 2001). Following transfer, the blot was labeled, rinsed in 2XSSC briefly and allowed to dry at ambient temperature. 10 mL hybridization solution was prepared containing 0.4 mL dH_2O, 0.5 mL 20% SDS, 3 mL 20XSSC, 5mL formamide, 1 mL 50X Denhardt's solution. Denature 0.1 mL sheared DNA [initial concentration: 10mg/mL] at 95°C X 5 min and add to prewarmed hybridization solution (42°C) immediately prior to use.

Membranes containing DNA were allowed to pre-hybridize for a minimum of 2 hours, after which pre-hybridization solution was removed and replaced with 10 mL fresh hybridization solution containing purified radiolabeled probe. Hybridization was typically allowed to proceed over night at 42°C. Following incubation with probe, washes were performed as follows: Rinse for 10-15 s in 200 mL 2XSSC at room temp; 200 mL 2XSSC/0.5% SDS at 25°C for 30 min; 200 mL 2XSSC/0.5% SDS at 60°C for 30 min; 200 mL 0.5XSSC/0.5% SDS at 60°C for 10 min. After washing, membranes were wrapped in saran-wrap and exposed to Amersham Hyperfilm ECL for 3-5 days at -80°C then developed.

Mouse colony establishment and maintenance

Chimeric mice were obtained first from the OHSU Transgenic Core facility. In all, 5 male chimeras were derived from two cell lines and kept in a pathogen free barrier room at OHSU. Ten female C57/BL6 mice were obtained from the Jackson Laboratory (Bar Harbor, ME) and delivered to the OHSU Department of Comparative Medicine (Protocol #A727). Two female mice were placed with each male chimera. Mice were monitored regularly for offspring. Individual mice from each litter of pups were ear-tagged, tail snips obtained and gender determined. High molecular weight genomic DNA was extracted from tail clips using Qiagen Genomic-tip 100/G columns and buffer set according to protocol (Qiagen, Cat#10243/19060). Genomic DNA derived from pups was then used to genotype each pup by Southern blotting according to methods described above. Later, pups were monitored and cataloged according to coat color only.

Offspring exhibiting black coat color were weaned from parents between 25 and 30 days of age, and non-essential mice were delivered to the Department of Animal care for additional use or euthanasia.

Rapid genotype screening

A unique feature of the targeting vector includes the ability of the *neo*-cassette to self-excise upon transmission through the male germline. PCR primers were designed to take advantage of this phenomenon for use in a rapid genotyping screen in agouti mice. Primer Genotype2.F was designed from mouse genomic sequence 5' of the point of insertion of the *neo*-cassette. Genotype2.F was paired with E1-2R (Appendix I) for use with Advantage Genomic PCR reagents (Clonetech Cat#K1906-1). A two-step PCR reaction using mouse genomic DNA was prepared as follows (1X reaction, all concentrations initial):

34.3 µL	dH_2O
5 µL	10X Reaction buffer
1 µL	50X dNTP mix [10 µM each]
2 µL	Forward Primer [25 µM]
2 µL	Reverse Primer [25 µM]
2.2 µL	Magnesium acetate [25 mM]
1 µL	50X Advantage Polymerase
2.5 µL	Genomic template (~ 100 ng/µL)
50 µL	total

Conditions:
95 °C X 1'
[95 °C X 30" / 68 °C X 1 min] X 35 cycle
68 °C X 3 min

Following amplification, PCR products were visualized by electrophoresis on a 0.8% agarose gel in 1X TAE and image capture following routine procedures.

Assessment of *Creld1* expression in heart tissue

A human cardiovascular system multiple tissue Northern blot was obtained from Clontech (Cat#7791-1). This Northern blot consists of approximately 2 μg poly A+ RNA from separate cardiovascular tissues including aorta, left and right atria, left and right ventricle as well as apex, total adult and total fetal heart tissue. The blot was probed using 50 ng gel-purified full-length human *CRELD1* cDNA labeled with [α^{32}P]dCTP as described above. Hybridization solution was 'ExpressHyb' supplied with the membrane. Probe was added directly to fresh ExpressHyb and allowed to hybridize for 1 h at 68°C. All washed were carried out according to kit protocol, followed by exposure to film as noted previously for 24-72 hours.

Yeast two-hybrid library

A pre-transformed human heart "Matchmaker" cDNA library was obtained from Clontech (Cat#HY4042AH). Briefly, the library was made from mRNA from normal, whole heart tissue pooled from three adult males. The number of independent clones was estimated by Clontech to be 3.5 X 10^6 with an average size cDNA insert of 2.0 kb. The "MatchMaker" Two-Hybrid System 3 screening components (including cloning and control vectors, yeast strains AH109 and Y187, carrier DNA, sequencing primers and

protocols) were also obtained from Clontech (Cat#K1612-1). This system is a GAL4-based system used in detecting protein-protein interactions.

Assembly of the DNA binding domain fusion construct

A strategy was designed to fuse amino acids 46 through 178 (nucleotides 137 to 530 of cDNA) of *CRELD1* to the GAL4 DNA-BD. This constitutes approximately 31% of the protein, beginning just downstream of the secretion signal sequence and ending in the center of the first of four predicted EGF-like domains. Two primers were designed to amplify the appropriate region of the *CRELD1* cDNA. These primers, named Y2H3F and Y2H2R (Appendix I), were used in PCR with Gibco BRL *Taq* DNA Polymerase and reagents. The reaction conditions were as follows (note: all concentrations are initial):

35 µL dH_2O
5 µL 10X Reaction buffer
2 µL dNTPs [10 mM]
2 µL Forward Primer Y2H3F [25 µM]
2 µL Reverse Primer Y2H2R [25 µM]
2 µL $MgCl_2$ [50 mM]
1 µL Taq DNA Polymerase
1 µL Template (full length cDNA in pcDNA3.1)
100 µL total

Conditions: 95 °C X 5'
[95 °C X 1' / 60 °C X 1' / 72 °C X 2'] X 30 cycles
72 °C X 10'

PCR product was gel purified then precipitated with 3M sodium acetate and 70% ethanol as described earlier. The fragment was then cloned into pMOS*Blue* blunt ended vector using reagents and methods included with the vector (Amersham Pharmacia Biotech).

Mini-preps were made from colonies according to standard methods and sequencing reactions performed using primer T7 to confirm proper sequence as well as assess orientation. Suitable clones were selected for further manipulation. A double restriction reaction was prepared using *BamHI* and *NdeI* (Boehringer Mannheim, Cat#220612 and #1040219 respectively) resulting in the excision of an insert with appropriate "sticky" ends. Vector pGBKT7(kan^r), the GAL4 DNA-BD fusion vector, was treated in an identical fashion. 20 µL reactions were made using 2 µg plasmid(s), 2µL Buffer B, 2µL each restriction enzyme and qs dH_2O; incubate at 37 °C for 2 h. Bands representing desired fragments were gel purified as described and standard ligation reactions were assembled using 1:4 ratio vector:insert. Transformants were selected on kanamycin (50µg/mL) plates. Plasmids were purified from resulting colonies and sequenced using vector T7(c) (Appendix I). Completed DNA-BD fusion plasmid was designated "pGBKT7.Y2H2".

Yeast phenotype testing

To make certain yeast stocks contained suitable host strains, phenotype was assessed based on the ability of each strain to grow on a variety of plates lacking essential nutrients. This was accomplished by streaking yeast strains AH109 (1n, MATa) and Y187 (1n, MATα) from frozen stocks on adenine-supplemented (complete) YPD plates (YPDA) for several days @ 30°C. A single colony was then used to inoculate SD plates lacking one of the following: adenine, methionine, tryptophan, leucine, histidine and uracil. AH109 is auxotrophic for listed amino acids except methionine; Y187 is

auxotrophic for all listed amino acids (Note: for all reagent stocks and recipes, refer to the Yeast Protocols Handbook ['YPH' from Clontech, PT3024-1] Appendix C: Media Recipes and Appendix D: Solution Formulations).

Yeast transformations and controls

Prior to screening the library, a series of control transformations, mating and phenotype testing was performed to reduce or eliminate the possibility of false positives and to verify the transformation markers of both bait and library plasmids. Using the small scale LiAc transformation procedure (YPH, Clontech), 100 ng of plasmids pCL1 (contains wild-type GAL4 as a positive control for the *βgal* assay), pGBKT7 (empty DNA-BD fusion plasmid), pGBKT7.Y2H2 ('bait' construct containing *CRELD1*), pGADT7 (empty DNA-AD fusion plasmid), DNA-BD/p53 (DNA-BD fusion with murine p53), pGADT7/T (AD-fusion with SV40 large T-antigen) were transformed into either the AH109 (in the case of DNA BD fusions or 'bait' plasmids) or Y189 (library plasmids) . For each transformation, 100 µL of 500 µL was plated. Additionally, serial dilutions of 1:10, 1:100 and 1:1000 were made and 100 µL each was plated. Transformations were grown on appropriate SD plates according to plasmid markers (SD/-Leu for DNA-AD vectors; SD/-Trp for DNA-BD vectors) and incubated 2-4 days at 30°C until colonies appeared. Colonies were counted and transformation efficiency calculated according to section V., YPH. Strain AH109 containing pGBKT7.Y2H2 was mated with strain Y189 containing pGADT7 (DNA-AD plasmid) and plated on quadruple drop-out media (QDO) to test the 'bait' plasmid for the ability to auto-activate

nutritional markers in the presence of the GAL4 activation domain. Later, these diploids were grown on SD -Leu/-Trp followed by *βgal* assay, again as a check for the ability of the bait plasmid to drive the reporter alone or interact with the empty DNA-AD plasmid. Further, strain Y189 harboring plasmid pGADT7/T was mated with AH109 containing DNA-BD-p53 and grown on QDO as a positive control, since SV40 large T-antigen and p53 are known to interact in the yeast two-hybrid assay (Li and Fields, 1993).

Yeast two-hybrid library screen

A 1.0 mL aliquot of the pre-transformed human heart cDNA library in yeast strain Y187 was mated with AH109 containing pGBKT7.Y2H2. This was accomplished by inoculating 50 mL SD/-Trp with a single 'bait' colony and incubating at 30°C while shaking until OD_{600} ~1.3. The cells were pelleted then resuspended in approximately 5 mL SD/-Trp. Next, the entire library was added to AH109/pGBKT7.Y2H2. The final volume was brought to 50 mL with YPDA/kan in a sterile flask shaking gently overnight (20-50 rpm, 30°C) to allow mating. Cells were then collected according to kit guidelines and resuspended in approximately 10 mL YPDA/kan. A portion of the mating mixture was reserved in order to titer the library and calculate mating efficiency. The latter was accomplished by spreading 100 µL of a 1:10, 1:100, 1:1000 and 1:10,000 dilution onto SD/-Leu, SD/-Trp and SD/-Leu/-Trp plates and incubating for 3-5 days at 30°C (until discreet colonies appeared). The remaining mating mixture was plated (approximately 200 µL per plate) onto 150 mm QDO plates and incubated for a minimum of 9 days at 30°C. All single, well isolated colonies from large library plates were transferred using a

sterile loop to 100 mm QDO plates placed over a grid. Plates were marked for orientation and individual colonies were given a binomial tag which included plate number and grid position (i.e. 14.17). These plates were incubated an additional 3-5 days prior to *βgal* assay.

β-Galactosidase assay

The *lacZ* reporter lies downstream of two separate GAL4 responsive upstream activating sequences in the diploid yeast. In order to qualitatively assess activation of this reporter, *βgal* assays were performed on all transferred yeast following the colony filter lift method (YPH Section VI.(C)). Briefly, sterile Whatman #5 filters were placed on top of yeast plates and gently rubbed with forceps to get colonies to adhere to the filter. Filters were pierced with a syringe and marked in ink for later orientation. The filter was then removed from the surface of the plate and plunged into N_2(l) for 10-15 s, followed by thawing at room temperature for several minutes in order to 'crack' the yeast. Filters were then placed colony side up on presoaked filter paper containing freshly prepared Z buffer/X-gal as described in the YPH. Filters were incubated at 30°C and monitored every 30 min (up to 8 hours) for color change.

Yeast plasmid recovery and analysis

Colonies showing rapid and robust color change in the *βgal* assays were selected for further analysis. Plasmids representing library clones of interest were recovered

using the YeastMaker yeast plasmid isolation kit (Clontech, Cat#K1611-1) and Chroma-spin 1000 columns following supplied protocol. After plasmid isolation, library inserts were PCR amplified using vector embedded primers and Advantage 2 PCR kit components as follows (note: all concentrations initial):

	36 µL dH_2O
	5 µL 10X Reaction buffer
	1 µL 50X dNTPs
	1 µL Forward Primer 5' AD-LD amplimer [20 µM]
	1 µL Reverse Primer 3' AD-LD amplimer [20 µM]
	1 µL 50X Adv. 2 Polymerase mix
	5 µL Template
	50 µL total
Conditions:	94 °C X 1'
	[94 °C X 30" / 68 °C X 3'] X 30 cycles
	68 °C X 3'

PCR products were visualized according to standard methods, followed by purification of each PCR reaction using Qiaquick PCR purification columns and a vacuum manifold according to manufacturers instructions (Qiagen Cat#28106). Purified PCR product was submitted to the Vollum Sequencing Core Laboratory for direct sequencing using 5' LD AD amplimer. Electronic sequence files were received and raw electropherograms were viewed using the Chromas (v.1.45) program. Successful reactions were manually edited to remove vector-based regions followed by comparison against the database at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/) using standard blastn parameters. Next, the sequence was analyzed using the DNA->Protein function (http://www.expasy.ch/tools/#translate) at Expert Protein Analysis System (ExPASy)

made available online by the Swiss Institute of Bioinformatics. Inserts were translated in-frame relative to the sequence of the upstream DNA Activation Domain of pACT2. Translated sequence was then compared against the NCBI database a second time using blastp. A list of all DNA insert and translated sequence was compiled. Comparisons were made between nucleotide and predicted proteins. Clones whose "blastn" result matched the "blastp" result were selected for additional study. Clones whose predicted nucleotide sequence diverged from the predicted translation were judged to be out of frame relative to the DNA-AD and thus assessment of this subset of clones was discontinued.

Confirmation of putative positive two-hybrid clones

Library clones selected for further analysis were confirmed for their *βgal* phenotype (as diploids) as described previously. Concentration of plasmids recovered from the diploid yeast were quantified by UV spectrophotometry, and approximately 50 ng each plasmid was used to transform DH5α using standard methods. Plasmids were amplified in 3 mL cultures and mini-preps performed. Confirmation of appropriate insert-containing plasmid was made by PCR using conditions noted above, followed by visualization on an agarose gel and comparing to original insert size. Furthermore, to be certain all plasmids contained expected library clones, each plasmid recovered from *E. coli* was sequenced and reassessed using methods and online tools described previously. Following these confirmation steps, a set of reciprocal experiments were carried out by moving each library plasmid from strain Y187 to AH109 and by placing bait and control

plasmids into strain Y187. To accomplish this, small scale yeast transformations were performed using each plasmid as described in YPH, Section V.(E). Included with the library inserts, empty vector pGADT7 was transformed into yeast strain AH109 as a control. Additionally, using identical methods, strain Y187 was transformed with pGBKT7.Lam, pGBKT7 alone, or pGBKT7.Y2H2. Haploid strains were then grown on SD/-Leu (AH109) or SD/-Trp (Y187) followed by *βgal* assay, again as a check for auto-activation. Each library insert in strain AH109 then underwent three crosses: one with empty pGBKT7, one with pGBKT7.Lam and a third with pGBKT7.Y2H2 (*CRELD1* bait plasmid). Additionally, pGADT7 (empty DNA-AD plasmid) in AH109 was mated with Y187 containing pGBKT7 alone or pGBKT7.Y2H2. Mated yeast were plated on SD/-Leu/-Trp and QDO, to confirm both phenotype and expectations of control crosses. After 3 to 5 days of growth at 30°C, *βgal* assays were performed to confirm earlier results.

***In vitro* transcription/translation**

Library plasmids containing inserts of interest, as well as bait plasmid pGBKT7.Y2H2 were used as templates in coupled transcription/translation reactions utilizing a TNT T7 Coupled Wheat Germ Extract System (Promega Cat#L4140). Reactions were prepared following manufacturer's guidelines (components were added in order listed, mixing gently after the addition of each reagent):

25 µL TNT wheat germ extract
2 µL TNT reaction buffer
1 µL TNT RNA polymerase (T7)
1 µL amino acid mixture minus Cysteine [1µM]

5	µL	^{35}S-Cysteine (Amersham Pharmacia Cat # SJ232)
1	µL	RNAsin Ribonuclease inhibitor (Promega Cat#N2111)
1	µg	DNA Template (circular plasimd)
		dH_2O quantity sufficient
50	µL	total

Incubate: 30°C x 90'

Following transcription/translation, 5, 10 and 20 µL aliquots were loaded onto a 12% acrylamide gel. Poly-acrylamide gel elctrophoresis was performed following standard procedures (Ausubel, 2001), after which the gel was washed in a solution of 7% methanol, 7% acetic acid and 1% glycerol for 5 minutes. The gel was then mounted on Whatman 3MM paper, covered with saran wrap and placed on a slab drier under vacuum at 80 °C for a minimum of 1 hour. Once dry, autoradiograms were obtained by exposing film to gels for 48-72 hours and developed following routine procedures.

Co-immunoprecipitation of labeled translation products

Co-immunoprecipitation experiments were carried out using newly synthesized *in vitro* transcription/translation protein products. This was accomplished by using the Clontech MATCHMAKER Co-IP kit (Cat# K1316-1), and was facilitated by virtue of the fact that the pGBKT7.Y2H2 insert was tagged with c-Myc while library inserts were tagged with the HA-epitope. Briefly, duplicate experiments were performed by mixing 10 µL each *in vitro* translated ^{35}S-cysteine-labeled reaction product and placing on a nutator at 4°C for 90 minutes. Kit supplied positive controls were used in parallel. Following this initial mixing step, either α-c-Myc or α–HA antibodies (1 µg) were added

to each reaction. Negative controls were performed by mixing c-Myc tagged translation products alone with α-HA antibody, and HA tagged translation products alone with α-c-Myc antibody. Each mixture was further incubated again on a nutator at 4°C for 90 minutes. Protein A beads (3 µL bed volume) were prepared by washing in 200 µL cold PBS followed by centrifugation at 7000 RPM x 30". Beads were brought to original volume by adding 3 µL PBS and the suspension was added to the reaction tube followed by mixing at 4°C for 90 minutes. Washes were performed by adding 100 µL kit-supplied Wash Buffers 1 and 2 and pelleting beads at 7000 RPM for 10" (5 X per buffer). SDS-PAGE loading buffer (20 µL) was added directly to the beads followed by heating (80°C for 5'). Samples were visualized by autoradiograms obtained following electrophoresis procedures as described above.

CRELD1 GST-Fusion

The *CRELD1* insert used in the yeast two-hybrid library screen was also expressed as a GST fusion protein after cloning into the pGEX-2T vector (Amersham-Pharmacia Cat#27-4801-01). A 20 ml starter culture of *E. coli* transformed with the fusion construct was grown in LB/Amp supplemented with 2% glucose at 37°C overnight in a shaking incubator. The following day, this culture was used to seed 1 liter of prewarmed LB/Amp and allowed to incubate at 30°C until the OD_{600} >1, at which point 0.4 mM IPTG was added followed by one additional hour of incubation. In order to extract the protein, cell pellets were collected (10' at 4000 RPM) and resuspended in cold PBS (1 ml per 4 g cells) containing protease inhibitor cocktail (Sigma Cat# P8465)

according to manufacturer's guidelines. Cells were then passed through a French press at least two times and the lysis was cleared at 16,000 RPM, 4°C X 30' in 50 ml polypropylene tubes using a super speed centrifuge . The cleared lysate was reserved, while the insoluble pellet was dissolved in 8M urea followed by dialysis in a buffer composed of 20 mM Tris pH 7.5, 1mM DTT, 1 mm EDTA and 250 mM NaCl. The dialysis of the solublized pellet was performed in 1liter of buffer at 4°C. Dialysis tubing containing proteins was moved to fresh buffer at least three times, allowing a minimum of 1 hour between transfers. The concentrated pellet as well as the cleared lysate were then added to 1 mL GST resin (Sigma Cat#G4510) followed by gentle shaking at 4°C for 30 minutes. GST resin was collected, washed 2X in excess PBS (approximately 20 ml PBS per ml resin) and samples were eluted from resin in 10 ml 25 mM glutathione (Sigma Cat#G4251). A second round of dialysis was performed on these supernatants as described above, after which 10 and 20 µL aliquots were visualized both by Western blot as well as Comassie stained protein gels. Electrophoresis and transfer to PVDF membranes were carried out according to standard methods (Ausubel, 2001). Western analysis was performed by incubating the membrane with mouse α-GST primary antibody (1:500; Santa Cruz Biotechnology Cat# SC138) for 1 hour followed by 3 X 15 ' washes in TBST. A second 1 hour incubation was performed using sheep α-mouse HRP conjugate (1:10,000; Amersham Pharmacia Cat#NA931V) as the secondary antibody followed by 3 X 15' washes in TBST. Finally, detection was performed using SuperSignal West Pico Chemiluminescent Substrate (Pierce Cat#34080T) according to manufacturers guidelines.

Chapter III

Results

Results

Library screen and isolation of genomic clone

In order go generate *Creld1* deficient mice, it was first necessary to clone the murine *Creld1* gene. Thus, a genomic library was screened to isolate a fragment of murine *Creld1* suitable to serve as a template for replacement cassette construction. This library, prepared by Eric P. Fox at the Transgenic Mouse Core Facility at Oregon Health Sciences University in 1998, was derived from mouse strain 129 murine embryonic stem (ES) cells. After obtaining an aliquot of this reagent, the library was titered and found to contain an average of 2.9×10^7 pfu/mL.

To generate a probe for use in screening the library, mouse EST sequence was first compared to the known human cDNA sequence. Primers were then chosen from areas of greatest similarity. These primers were used in PCR reactions to amplify regions of mouse genomic DNA, and the resulting fragments were cloned and sequenced. A PCR product of 538 bp was obtained in this fashion using primers E1-2F and E1-2R. Sequence data demonstrated this PCR product contained a portion of murine *Creld1* exon 1, intron 1 in its entirety and a portion of exon 2. Closer examination revealed significant sequence identity between human and mouse *Creld1* exons 1 and 2 (91% and 87% respectively) though the sequence of the first intron was highly divergent between species. This PCR product was selected for use as a probe with which to screen the mouse genomic library.

A total of approximately 2×10^6 plaques were screened. Clones identified during the primary screen were confirmed positive by successive rounds of plaque

purification followed by re-plating, filter lifts and re-hybridization with the probe until all plaques on a given plate gave a positive signal. This process was repeated a minimum of three times per positive plaque. Viral DNA was obtained from 100 mL cultures inoculated with phage eluate using the Wizard Lamda Prep DNA Purification system following kit protocol. Though recovery efficiency varied, adequate amounts (ranging from 24 to 86 μg total) of viral DNA was obtained. Bacteriophage DNA was assessed by PCR using primers E1-2F and E1-2R, reasoning that those templates yielding positive results would contain at minimum two exons of the target gene and would be of greatest utility in terms of constructing the replacement cassette. Two individual clones were positive by this PCR assay. Of these, a single clone was selected for further characterization.

After recovering the murine insert from the bacteriophage vector, the fragment was subcloned into pBluescript II SK(-). Because the library was prepared by cloning genomic fragments into LamdaFix II using *SalI*, and because the insert was recovered using *NotI*, the subcloned fragment carried T7 and T3 ends derived from the multiple cloning sites of LamdaFix II (Appendix II). The subcloned library insert was thus used as the starting material for partial characterization of the mouse genomic fragment by restriction mapping. Oligonucleotide primers were end-labeled with T4 polynucleotide kinase and free nucleotide removed. These probes were then used on Southern blots harboring restricted insert. By using a combination of oligonucleotides known to localize to different portions of the complete subcloned genomic insert, a partial restriction map was deduced (figure 3).

Figure 3

A partial restriction map of the murine *Creld1* lamda clone. Shown at the top is the full length clone in 1 kb intervals (approximate). T7 was determined to be at the 5' end of the insert relative the first exon of the gene, and T3 at the 3' end. Informative enzymes are listed in the column on the left, and each individual enzyme pattern is illustrated. Solid lines indicate intact fragments, and approximate size of each fragment is shown immediately above. Dashed lines indicate regions of multiple recognition sites (for the given enzyme) and thus remain unresolved. The double bar above the single SacI site represents the probe used to screen the library, and includes portions of exons 1 and 2 (not shown).

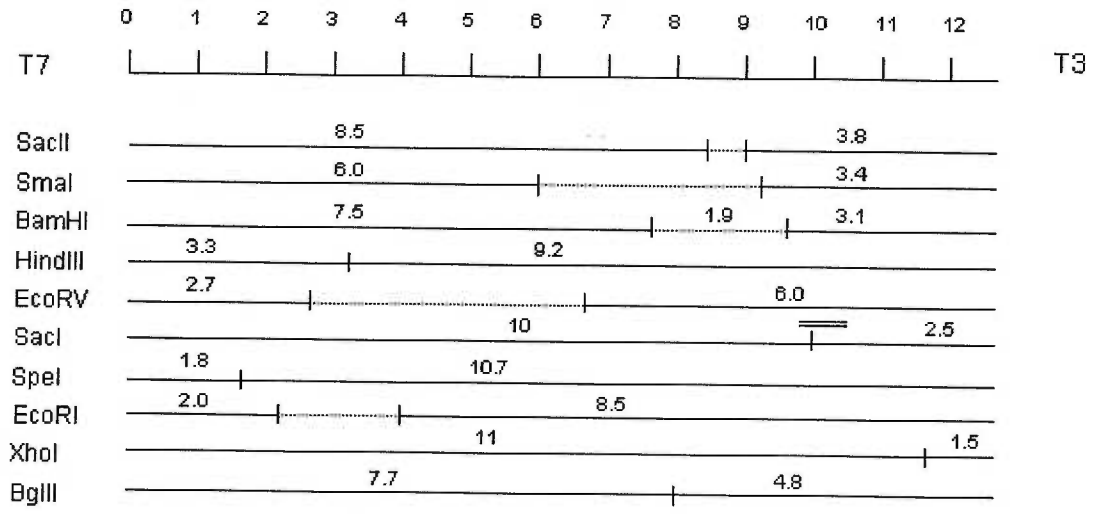

0 1 2 3 4 5 6 7 8 9 10 11 12
T7
T3
SacII
SmaI
BamHI
HindIII
EcoRV
SacI
SpeI
EcoRI
XhoI
BglII
8.5
3.8
6.0
3.4
7.5
1.9
3.1
3.3
9.2
2.7
6.0
10
2.5
1.8
10.7
2.0
8.5
11
1.5
7.7
4.8

Figure 3

Radiation hybrid panel

The mouse *Creld1* gene was localized within the mouse genome using the T31 Mouse Radiation Hybrid Panel RH04.02. This was done to test the prediction that the murine *Creld1* locus had been successfully cloned and that the gene mapped to mouse chromosome 6, an area known to by syntenic with distal human chromosome 3p (Blake et al., 2001). 100 cell lines were included in the analysis, and judged positive when a single unambiguous 528 bp band was visualized upon electrophoresis. All procedures were performed in duplicate and analyzed using the Jackson Laboratory t31 Mouse Radiation Hybrid Database. *Creld1* was localized to a 26.6 cR interval between D6Mit10 and D6Mit329, with highest anchor LOD score of 12.5 to D6Mit329, indicating the gene maps to distal (49.5 cM) chromosome 6. As shown in figure 4, this finding is in close agreement with predictions made from the human mapping data, which places *CRELD1* on chromosome 3p25.

Figure 4

Comparison of loci found on human chromosome 3p25 (on right) and distal mouse chromosome 6 (on left). Anchor loci from radiation hybrid analysis are shown in bold. Note the high degree of synteny between the two regions. The interval spans approximately 8 Mb physical distance in humans. The physical distance in mouse has not yet been determined. CAMK1, Calcium/Calmodulin dependent protein kinase I; VHL, von Hippel Lindau Syndrome; HRH-1, Histamine Receptor H1; TIMP4, Tissue inhibitor of Metalloproteinase 4.

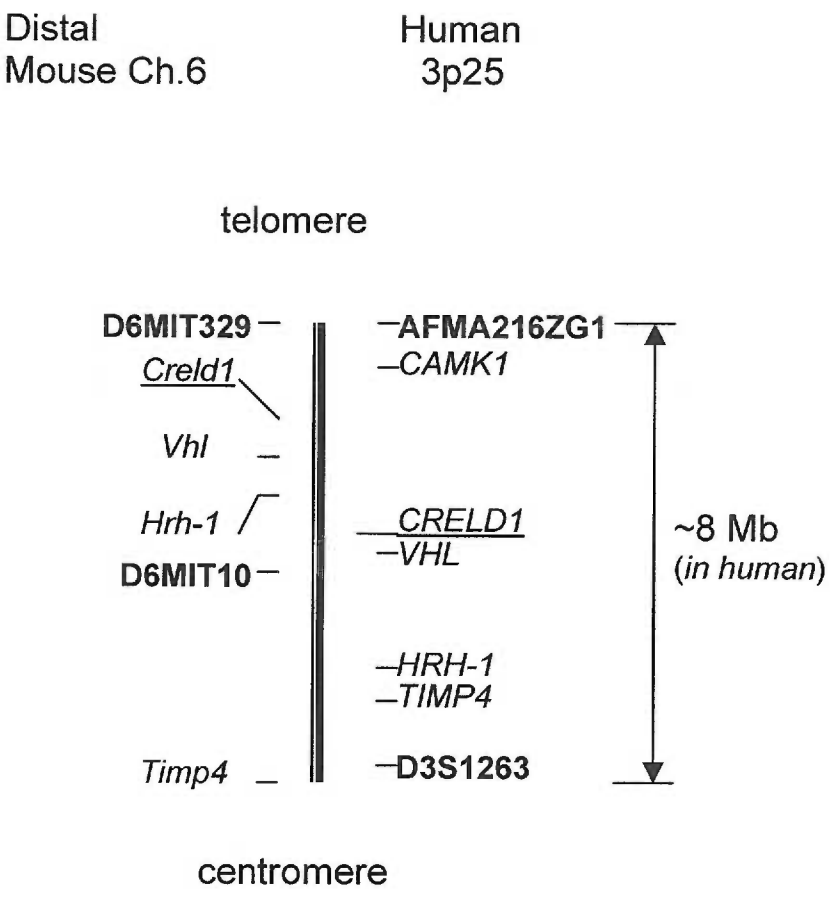

Distal
Mouse Ch.6
Human
3p25
telomere
D6MIT329
Creld1
Vhl
Hrh-1
D6MIT10
Timp4
AFMA216ZG1
CAMK1
CRELD1
VHL
HRH-1
TIMP4
D3S1263
~8 Mb
(in human)
centromere

Figure 4

Complete characterization of the mouse *Creld1* cDNA

A mouse EST (GenBank accession#BE200505) was obtained and sequenced using vector primers T7 (5') and T3 (3'). Insert specific primers (M.Ex3-4F, M.Ex5-6F and M.Ex8-9F) were also used. When compared to the human cDNA, the murine orthologous cDNA was shown to be 89% identical at the nucleotide level. The translated sequence was found to be 91% identical (94% similar) between mouse and human (Appendix II). The mouse sequence and mapping data were deposited at the Mouse Genome Database (http://www.informatics.jax.org/nomen/), and the mouse locus was given the official designation *Creld1* (J72511).

***Creld1* knockout targeting vector**

A replacement vector was assembled for use in targeted deletion of *Creld1* via homologous recombination in mouse ES cells (figure 5). Assembly details of this vector are summarized in material and methods. The completed construct contains a self-excising neomycin resistance cassette (Bunting et al., 1999) termed ACN. This cassette contains a *neo*-gene downstream of the mouse RNA Polymerase II promotor, which itself is downstream of a gene encoding the *Cre* recombinase. The *Cre* gene in turn is under the control of the murine testis-specific angiotensin-converting enzyme promotor, tACE, which is activated during spermatogenesis. This entire sequence is flanked by two *loxP* sites. When ES cells carrying a targeted allele contribute to the germline of male mice, *Cre* expression occurs during gametogenesis causing Cre-mediated recombination between the *loxP* sites thereby excising the entire cassette. The resulting paternal

Figure 5

Visualization of the strategy for assembling the completed targeting vector. The ACN contains a neomycin resistance gene downstream of mouse RNA PolII promotor. The Cre-recombinase is upstream of *neo* and is under control of the murine angiotensin converting enzyme (tACE) promotor. The entire cassette was inserted into the BglII site of BamHI/BglII/SacI linker (shown in yellow). The linker had been positioned in between native BamHI and SacI sites after the removal of approximately 2 kb of the cirrin gene and 5'UTR. Downstream of the 3' homologous arm, a thymidine kinase (TK, shown in red) cassette was positioned to aid in negative selection. The entire targeting vector was assembled on a pEGFP vector backbone after excision of GFP coding sequence. Exons 1 and 2 of *CRELD1* are shown.

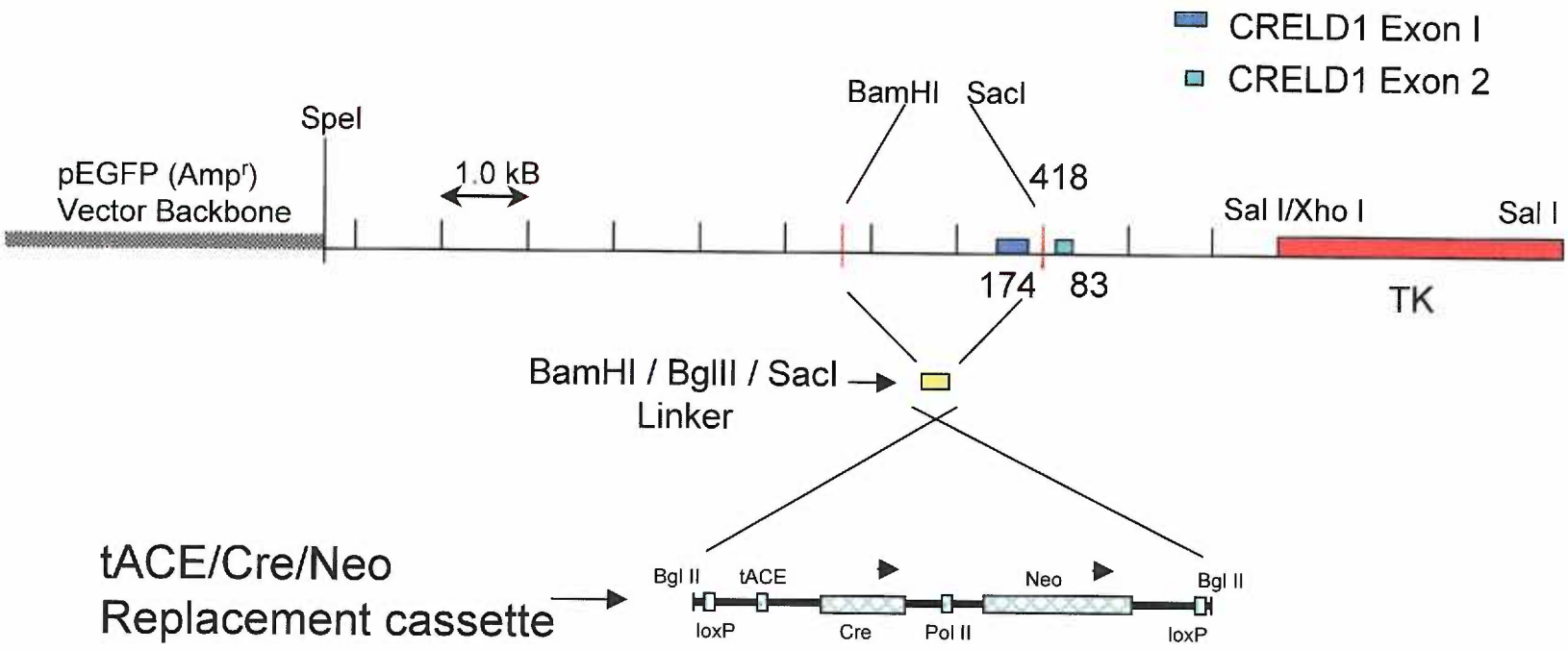

CRELD1 Exon I
CRELD1 Exon 2
BamHI
SacI
SpeI
pEGFP (Ampr)
Vector Backbone
1.0 kB
418
Sal I/Xho I
Sal I
174
83
TK
BamHI / BglII / SacI
Linker
tACE/Cre/Neo
Replacement cassette
Bgl II
tACE
Neo
Bgl II
loxP
Cre
Pol II
loxP

Figure 5

contribution to the offspring is either a wild-type or null allele. The benefit of this approach is that the null allele lacks the strong promotor typically used to drive *neo* expression, thus eliminating the possibility of phenotypic consequences resulting from the improper activation of nearby genes. It has been demonstrated that such effects can lead to ambiguous results in mouse models (Olson et al., 1996).

The ACN is flanked by two regions homologous to the *Creld1* locus. The upstream (5' relative to the *Creld1* ATG codon) arm of the finished replacement cassette spans approximately 6.3 kb. The region downstream of ACN harbors nearly 3 kb with homology to murine *Creld1*. The ACN-cassette itself was positioned within the genomic fragment such that all of exon 1, a portion of intron 1, and 1.9 kb of 5' UTR was eliminated (figure 5). This construct was delivered to the OHSU Transgenic / Gene Targeting Facility as well as to the Gene Targeted Mouse Service at the University of Cincinnati for ES cell culture, electroporation and blastocyst injection with the goal of obtaining chimeras for colony establishment.

Upon delivery of the targeting vector to the core laboratories, an aliquot of mouse strain 129 genomic DNA was used in the development of a Southern assay for later application in the positive identification of homologous recombination. Using a 1.8 kb portion of the genomic clone not included in the targeting vector, probes were designed and tested on total genomic DNA treated with a variety of restriction enzymes in pilot Southern blot experiments, the results of which are summarized in Appendix II, table I. By this method it was revealed that the probes were inadequate for the intended application. Further analysis of this 1.8 kb region demonstrated the presence of several

SINE/B1 and LINE elements as recognized by the RepeatMasker online sequence utility at (http://ftp.genome.washington.edu/RM/RepeatMasker.html) (figure 6).

Figure 6

(A) Diagram of the genomic clone, showing the positions of exons 1 and 2, as well as the 1.8 kb fragment at the 5' end of the sequence. (B) A detailed view of the subcloned 1.8 kb fragment. Three probes (1.8, N/S and B) are indicated by double bars. Sequencing primers (T7,T3, Frag1.8-2.seq and Frag1.8-3.seq) are shown. Also shown are repeat elements (hatched boxes) present within the sequence. Subtype of repeat and size in bp are indicated.

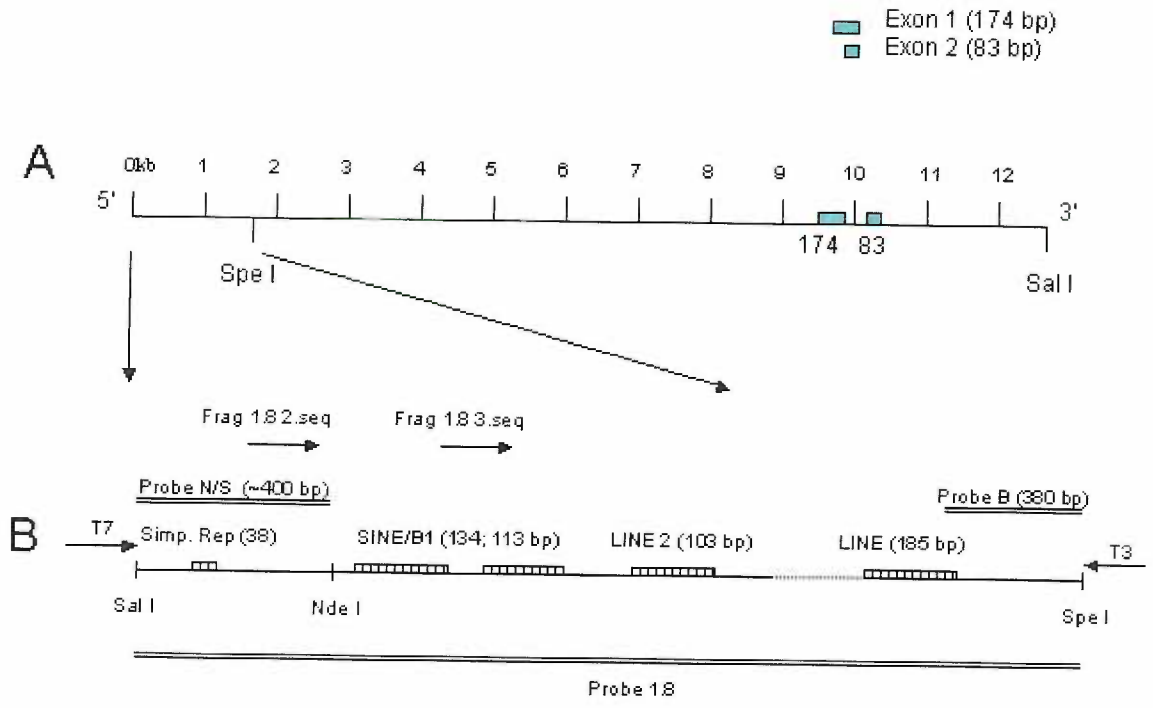

Exon 1 (174 bp)
Exon 2 (83 bp)
A
0kb
1
2
3
4
5
6
7
8
9
10
11
12
5'
3'
174
83
Spe I
Sal I
Frag 1.8.2.seq
Frag 1.8.3.seq
Probe N/S (~400 bp)
Probe B (380 bp)
B
T7
Simp. Rep (38)
SINE/B1 (134; 113 bp)
LINE 2 (103 bp)
LINE (185 bp)
T3
Sal I
Nde I
Spe I
Probe 1.8

Figure 6

Analysis of mouse chromosome 6 BACs

Three mouse BACs were obtained from Incyte Genomics. These BACs were identified using the Incyte BAC Mouse II PCR library screening service. Primers E1-2F and E1-2R were submitted to Incyte Genomics for the purposes of positively identifying BACs containing the region of interest by PCR. BACs were cultured according to the guidelines recommended by Incyte Genomics and these relatively large DNA molecules were isolated using the Qiagen Large Construct Kit. DNA was collected from each of three BACs, ranging from 4 to 8 μg total. After DNA extraction, PCR results were confirmed using the same primer set and conditions for each BAC. Partial characterization of each BAC was accomplished by using additional primers, designed from sequence obtained by scanning the EST database or designed directly from sequence generated from the genomic clone used in making the targeting vector. PCR products were produced and sequenced. One of the primer pairs (E2-3F.2 and E2-3R) encompassed the end of the clone obtained from the bacteriophage library, thus enabling the assembly and partial characterization of a contig extending into the *Creld1* gene downstream of the genomic clone. This was accomplished by using overlapping primer sets as described in materials and methods. Furthermore, restriction sites were identified from sequenced regions using Webcutter 2.0. Finally, several restriction digests were performed to further characterize the region. A summary of the position of these overlapping PCR products and restriction sites is illustrated in figure 7.

Figure 7

(A) Wild type *Creld1* locus. The ATG codon, exons 1 through 4 and critical restriction sites are labeled. Shown above this locus are PCR fragments, labeled with the primers used to amplify each fragment as well as relative size and location of each PCR product. For clarity, primer M.Ex3-4F (used with M.Ex5-6R) has been omitted. The probe used in genotyping both ES cells and mice is shown as a double bar below the locus. (B) Targeting vector shown in relation to the wild-type allele. (C) The *Creld1* locus shown after a homologous recombination event. Note the introduction of a BglII site downstream of the *neo*-cassette. The wild-type and mutant alleles, after restriction with BglII, are illustrated below the modified *Creld1* locus.

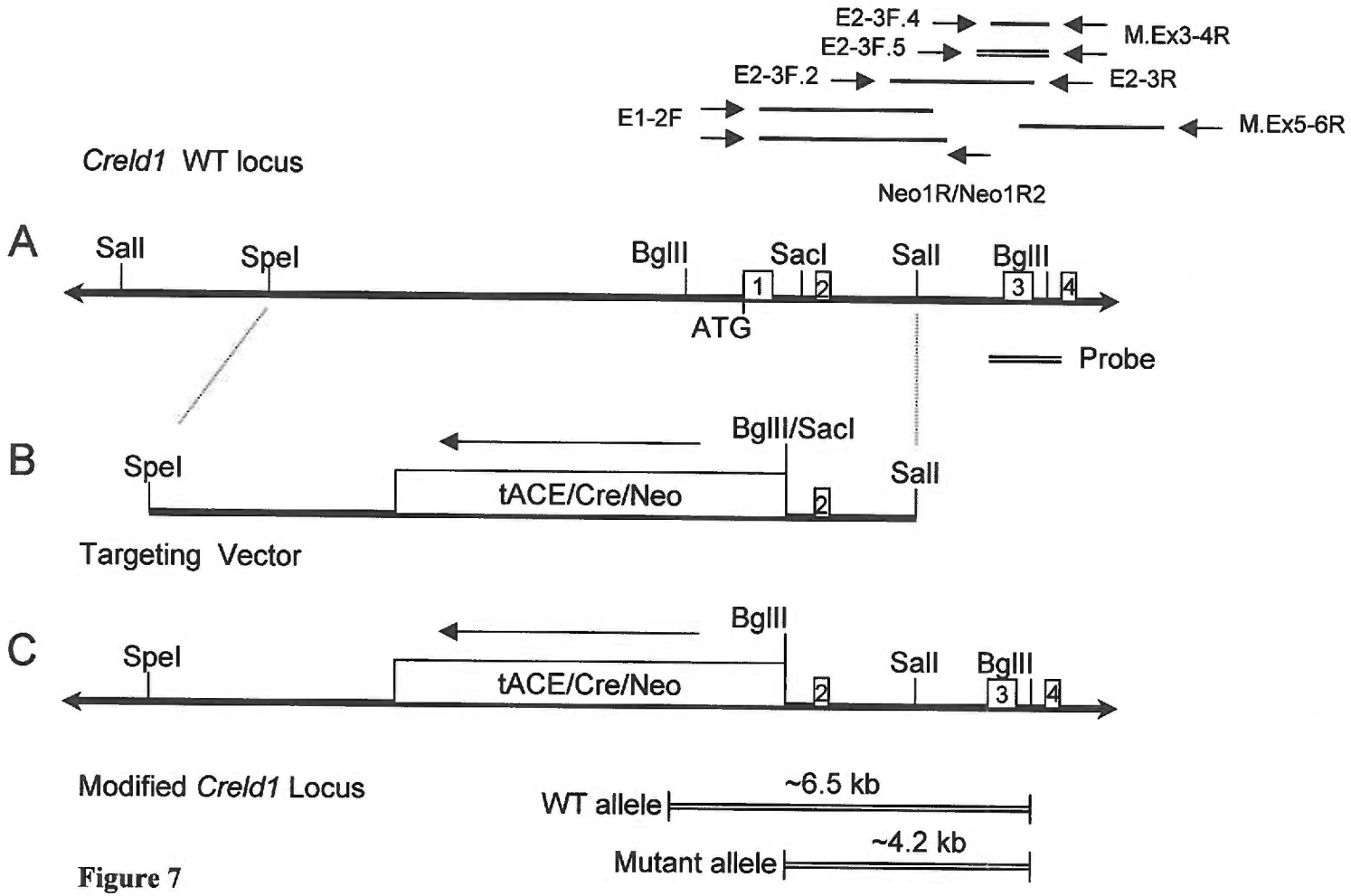

E2-3F.4
M.Ex3-4R
E2-3F.5
E2-3F.2
E2-3R
E1-2F
M.Ex5-6R
Neo1R/Neo1R2
Creld1 WT locus
A
SalI
SpeI
BglII
SacI
SalI
BglII
1
2
3
4
ATG
Probe
B
BglII/SacI
SpeI
tACE/Cre/Neo
SalI
2
Targeting Vector
C
BglII
SpeI
SalI
BglII
tACE/Cre/Neo
2
3
4
Modified Creld1 Locus
~6.5 kb
WT allele
~4.2 kb
Mutant allele

Figure 7

Using the sequence available from the BACs, a new probe for use in genotyping was derived by PCR. This probe was the result of amplification using primers E2-3F.5 and M.Ex3-4R (figure 7). Additional Southern blots were prepared. When high molecular weight genomic DNA was restricted using *BglII*, this probe annealed to a wild-type allele migrating at approximately 6.5 kb relative to size markers. Since the *neo*-cassette was inserted using *BglII*, it was evident that after homologous recombination, a targeted allele would yield a fragment of approximately 4.2 kb that would be recognized using the same probe (figure 7), thereby allowing discrimination between native and mutant alleles.

Analysis of ES cell genomic DNA

After receiving ES cell pellets (either in 96 well plates or 1.5 mL microfuge tubes), high molecular weight genomic DNA was obtained as described above. This DNA was treated with *BglII* and assessed by Southern blot. A total of 74 cell lines received from the OHSU facility were evaluated by this method. An additional 274 cell pellets delivered from the University of Cincinnati were treated in an identical fashion. Of the initial 74 cell lines screened from the OHSU core, 5 positives were identified and confirmed. This translates to a targeting efficiency of 6.8%. Similarly, of 274 cell lines screened from the University of Cincinnati facility, 15 positive clones (5.5%) were found (figure 8).

Figure 8

(A) Identification of a single recombinant ES cell line by Southern blot. Lane marked with (*); mutant band identified with an arrow; wild-type band shown above (present in each lane). (B) Southern blot confirming homologous recombination in 8 separate ES cell lines. M, HyperLadder molecular weight standard, reference bands (2.5,4 and 6 kb) labeled on the right. Four non-recombinant (negative control) cell lines labeled as "wild-type" shown (cell lines obtained from University of Cincinnati).

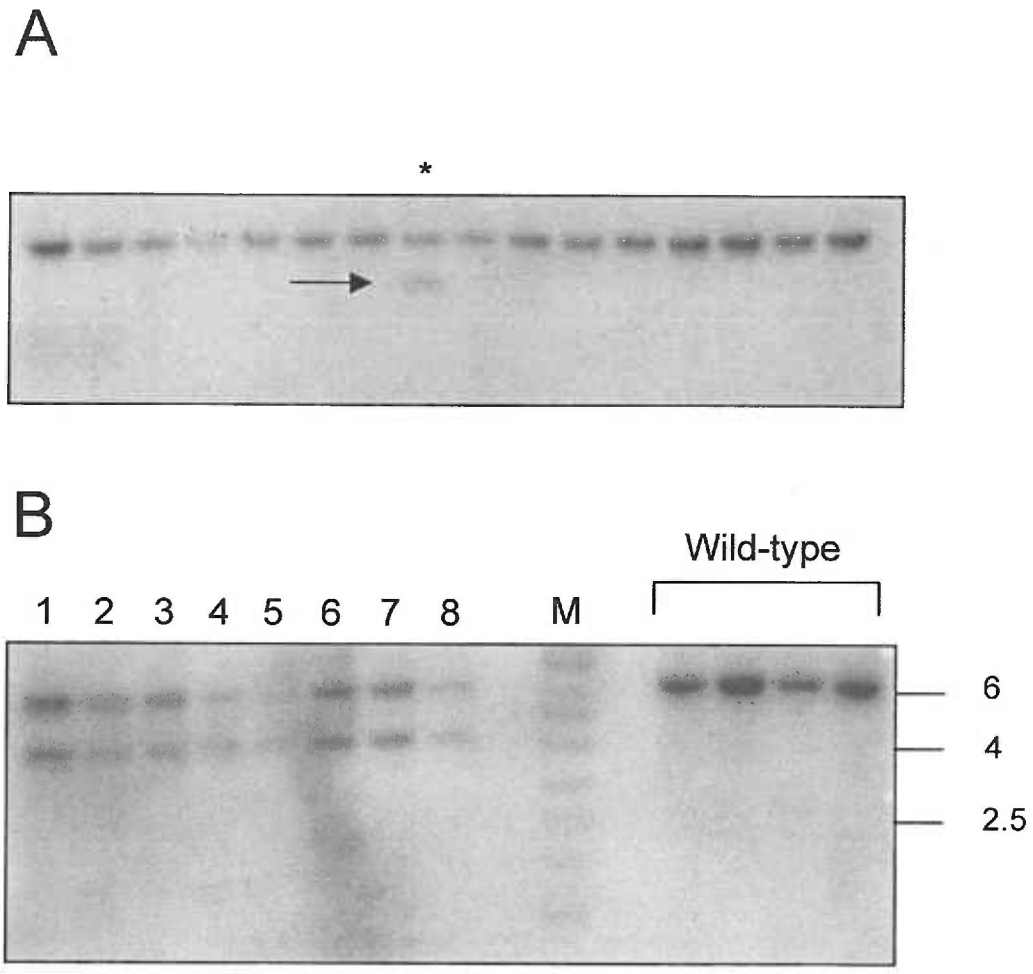


Figure 8

Establishment of mouse colony and breeding

Chimeric male mice were obtained from the OHSU transgenic facility. These male mice (5 total) were designated A, B, C, D and E. Each chimera was paired with two female C57/BL6 mice. Mice were monitored routinely for a period of 11 months. During this time, multiple litters were obtained from all male chimeras with the exception of mouse E for which litters were not obtained even after being paired for several months with fertile female mice. It was thus determined that mouse E was sterile and breeding was discontinued.

Initially, pups obtained from each fertile male were tagged using ear clips and logged into a notebook. At 3-4 weeks of age, a 1 cm sample of tail was removed after which the litter was weaned. The tail biopsy was used to extract high molecular weight genomic DNA. In this manner, DNA from the first 64 pups (representing at least 1 litter from each chimera) was obtained and assessed by Southern blot as described. Primarily, this was done to establish familiarity with these techniques since there were no pups showing agouti coat color on visual inspection. In all cases Southern assay results confirmed the visual observations. Weaned pups were donated to the Department of Animal Care (DAC) for euthanasia or for use in training exercises. Additional offspring were evaluated by visual inspection only prior to weaning and delivery to DAC. In order to assure the presence (or absence) of germ line transmission, a minimum of 50 pups were obtained from each chimera. As of this writing, a total of 229 pups have been produced. The individual totals are summarized in Table (B):

Table B: Offspring produced by OHSU chimeras

CHIMERA	PUPS (#)
A	57 Pups
B	58 Pups
C	52 Pups
D	62 Pups
Total	229 Pups

Because of the apparent lack of germ line transmission from any of our five original chimeras, additional chimeras were generated at the University of Cincinnati. Three rounds of blastocyst injections were performed with four independent ES cell lines. Pups were obtained from these blastocyst injections. A total of 16 pups were born, 6 of which were chimeric males. Each male exhibited a high proportion of agouti coat color, thus increasing the probability that the injected ES cells contributed to the reproductive organs in these mice. Thus far, 37 agouti pups have been derived from four of six chimeras, demonstrating a contribution of targeted ES cells to the germ line of each of the four chimeras. Further, these chimeras were derived from two separate ES cell lines. DNA recovered from tail biopsy from each pup was analyzed using the rapid PCR based assay detailed above. Of the initial 37 agouti pups received, 5 male and 4 female heterozygous pups were identified (figure 9). These pups were used to establish a *Creld1* knock-out colony at OHSU.

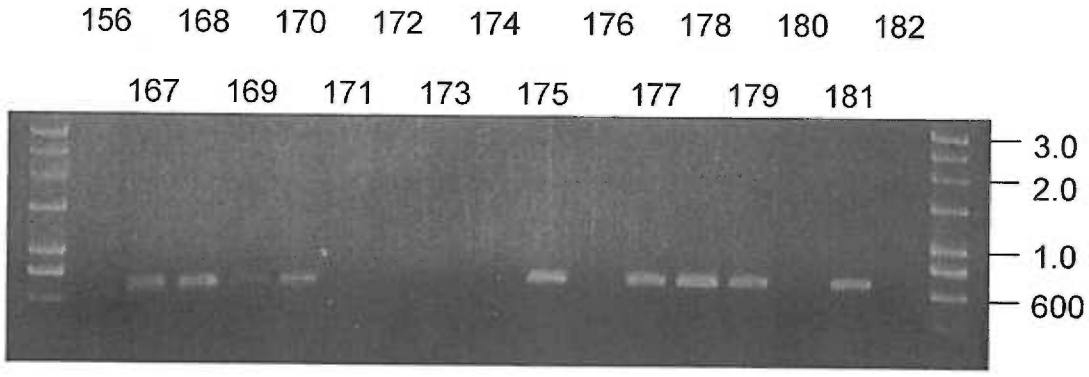


Figure 9

Agarose gel stained with ethidium bromide showing the presence (or absence) of a 720 bp product which detects the self-excision of the *neo*-cassette in heterozygous agouti pups. Presence of PCR product indicates an appropriately targeted insertion of the *neo*-cassette. Lack of PCR product indicates a wild type locus. Molecular weight of visible marker bands are indicated (HyperLadder).

CRELD1 homozygous null mice expire *in utero*

Initially, each of five male *Creld1* heterozygote-null mice were paired with two female C57/BL6 mice . Offspring from these breeding pairs were collected and genotyped both by PCR and Southern blot as previously outlined. Wild-type (+/+) and heterozygous (+/-) *Creld1* mice were obtained in essentially equal numbers as expected for the F_1 generation (Table C).

Table C: *Creld1* genotype distribution in F_1 and F_2 generations. Embryonic genotypes were determined following timed matings between e9.5-11.5.

	F_1	F_2	F_2 EMBRYOS
Wild Type (+/+)	29	31	14
Heterozygous (+/-)	30	52	35
Homozygous (-/-)	--	0	12

Next, a second round of crosses were performed between F_1 heterozygotes. DNA samples from the F_2 progeny were obtained and used to genotype each mouse. Over 80 F_2 mice were collected and analyzed. Of these, no homozygous *Creld1* pups were identified, suggesting that *Creld1* null mutants were dying prior to term. In order to confirm this hypothesis, timed matings were performed using heterozygous-null pairs. Female mice were sacrificed and embryos were harvested between e 9.5 - e11.5. A gross anatomical inspection of each embryo was made prior to DNA extraction and analysis. During these inspections a population of strikingly underdeveloped embryos were observed (Figure 10). Upon manipulation, the tissues of these smaller embryos were found to be qualitatively more fragile than those of larger, more fully formed littermates. Genotyping assays were performed using DNA extracted from embryos after dissection from placental tissues. As shown in Table C, embryonic F_2 genotypes followed the predicted Mendelian distribution, thus confirming the hypothesis that homozygous *Creld1*-null mutations are not compatible with survival to term.

Figure 10

A Creld1 null embryo (approximately embryonic day 11.5; top of figure) is shown compared to its wild-type litter mate. The mutant embryo shows a dramatic reduction in size. This qualitative size reduction is marked by severely underdeveloped limbs and a small mis-shapen head.

Figure 10

CRELD1 expression in the heart

We have hypothesized that *CRELD1* is involved in heart development. It has been shown previously by *in situ* hybridization that *CRELD1* is expressed in the myocardium and endocardial cushions in the developing heart (Rupp, Fouad et al., 2002). In order to assess *CRELD1* expression in individual cardiac tissues, a human cardiovascular system multiple tissue Northern blot was obtained and probed using full-length *CRELD1* cDNA as described. As shown in Figure 11, the 2.1 kb transcript is highly expressed in each tissue included on the blot. Additionally, there is evidence that the full-length cDNA recognizes a second slightly larger transcript as demonstrated by the presence of a weakly detectable band just above the primary transcript. Whether this is an alternatively spliced transcript or the transcript of *CRELD2* remains to be determined.

Figure 11

A human cardiovascular system multiple tissue Northern blot probed with full- length *CRELD1* cDNA. The major 2.1 kb transcript is apparent in all tissues examined. The presence of a second alternatively spliced (or closely related gene transcript) appears at approximately 2.8 kb and is particularly evident in the apex and aorta. Size markers are shown on the right. Lanes: RV, right ventricle, LV, left ventricle, RA, right atrium, LA, left atrium, APX, apex of heart, AO, aorta, AH, adult heart (total), FH, fetal heart (total). A β-actin control probe (to normalize loading) is shown below and appears on the blot as a doublet. β-actin is known to be over expressed in aorta.

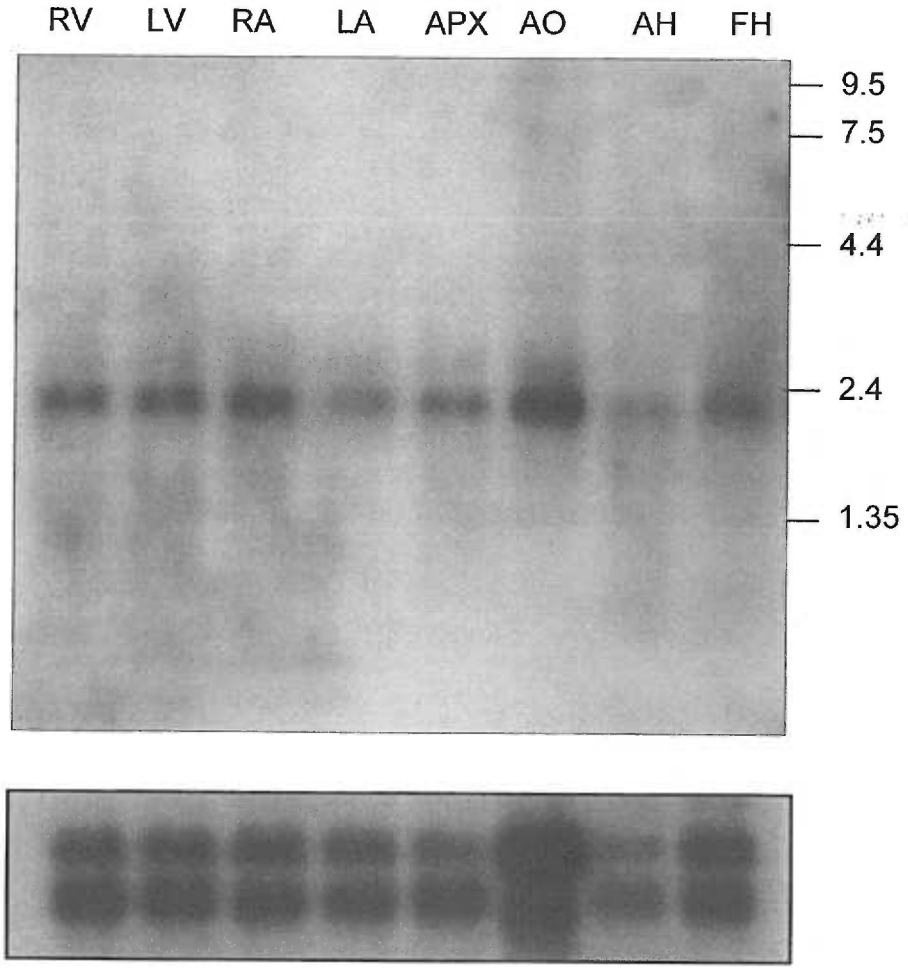


Figure 11

Yeast-two hybrid construct

In order to screen a pre-transformed human heart 'Matchmaker' cDNA library, a GAL4 DNA-BD fusion construct was assembled in pGBKT7. The CRELD1 protein is modular, composed of sets of tandem EGF-like and calcium binding EGF-like domains as illustrated in Figure 12. Additionally, two predicted transmembrane spanning domains lie at the carboxyl-terminal end of the protein, whereas at the amino-terminus, a secretion signal peptide is followed by a tryptophan/glutamic acid-rich region (WE domain). Using CRELD1 amino acid sequence to scan the database revealed proteins from diverse families containing each of these 'modules' with the exception of the Trp/Glu rich region. Thus, we hypothesized that this WE domain may be performing a unique function and that using this region of the protein in a two-hybrid library screen might provide insight into the biological role of CRELD1. Thus, the 'bait' construct contains a 132 amino acid fragment of CRELD1 fused downstream of the GAL4 DNA-BD. This peptide lacks the native CRELD1 secretion signal but encompasses nearly one third of the protein, including the tryptophan/glutamic acid-rich region (Figure 12).

Yeast transformation and library screen

Prior to transforming yeast with the bait and control plasmids, the phenotypes of strains AH109 and Y187 were confirmed as described in Materials and Methods. After verification of each phenotype, transformations were performed with test and control plasmids as listed previously. Haploid yeast were then mated and tested on either –Leu/-Trp plates (to ensure both DNA-BD and DNA-AD plasmids are maintained) or QDO

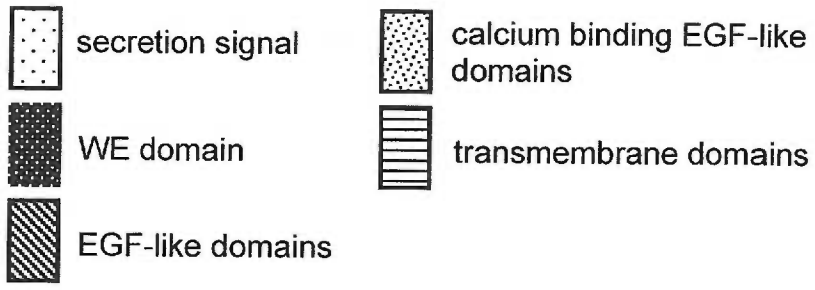
plates to assess nutritional marker activation and *βgal* assay reporter. In particular, diploid AH109/Y187 containing pGBKT7.Y2H2 (bait) and pGADT7 (empty AD vector) did not grow on QDO plates indicating the bait construct does not auto-activate nutritional markers and was suitable for use in the screen. All controls yielded expected outcomes with the exception of pCL1 (which encodes the full-length wild-type GAL4 protein and acts as a positive control for the *βgal* assay) for which no transformants were obtained. Nevertheless, the presence of both positive (blue) and negative (white) colonies on separate *βgal* assays with other test and library plasmids confirmed the assay was working. Transformation efficiency was relatively low, calculated to be between 100 to 500 cfu/µg plasmid for pGADT7 and pGBKT7.Y2H2 respectively.

Once appropriate transformants had been identified, an aliquot of the library in strain Y189 was screened by mating with AH109 containing pGBKT7.Y2H2 as outlined above. Following mating, serial dilutions were made and both titer and mating efficiency were calculated. The titer of the library itself was determined to be 3×10^7 cfu/mL; mating efficiency was calculated to be 1.8% and approximately 3.67×10^5 clones were screened.

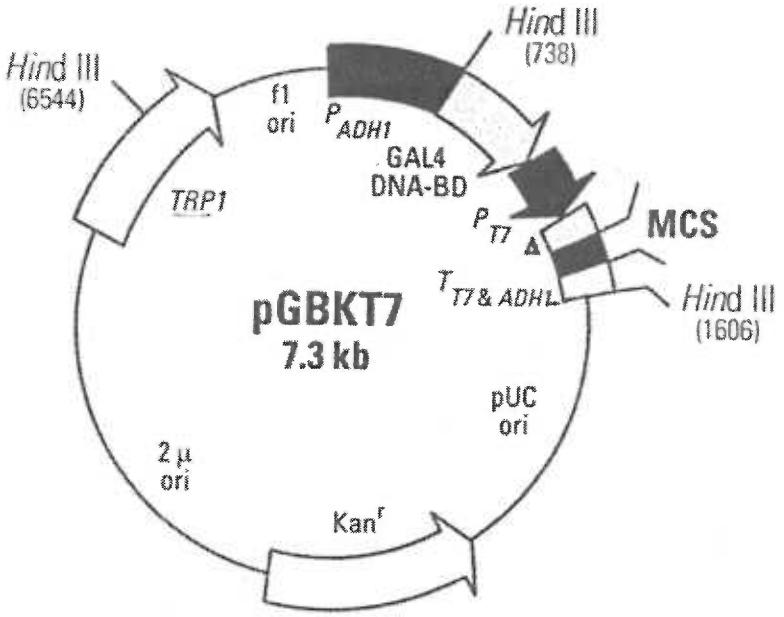
1137 individual colonies were transferred from library plates to 100mm QDO plates arranged on a grid. Filter lifts were made of these colonies and *βgal* assays performed. Colonies typically began to change color within 30-45 minutes, though the reaction was allowed to proceed for up to 8 hours. Colonies were chosen for further analysis based on robust color change within the first 2-4 hours. In this fashion, a subset of 80 colonies was selected and plasmids were obtained from yeast grown in liquid

Figure 12

A diagram of the *CRELD1* domain structure is shown. The WE (tryptophan-glutamic acid) domain (underlined) was cloned into the MCS of pGBKT7 (BamHI/NdeI), fusing it downstream of the GAL4 DNA-BD. This construct was then used to screen a human heart 'Matchmaker' cDNA library.



Subcloned fragment (cDNA nucleotides 137 – 530)



▲ c-Myc epitope tag

Clontech

Figure 12

culture. PCR products from each plasmid were produced and directly sequenced. As noted above, a strategy was chosen for further selection of positive clones based on sequence analysis. First, the nucleotide sequence was subjected to a blastn search, followed by translation (in frame relative to the GAL4 DNA-BD) and blastp search. Only those clones whose nucleotide sequence and predicted translation product were found to match one another were considered further. In this manner the final number of colonies was reduced to nine (Table D).

Given the prediction that CRELD1 is a membrane associated protein and the broad expression pattern evident on multiple tissue Northern blot and *in situ* hybridization those clones above showing restricted expression were eliminated from immediate consideration. Furthermore, clones encoding proteins whose cellular localization is primarily cytoplasmic or nuclear were also removed from the list of potential CRELD1 interaction candidates. Based largely on these criteria, the primary conclusion drawn from these studies is that CRELD1 is a potential binding partner for the extracellular protein decorin and the matricellular protein SPARC/Osteonectin/BM-40.

Yeast two-hybrid confirmation

The two putative positive clones identified in the library screen were re-tested by recovering each plasmid and transforming the opposing strains of yeast, followed by mating both with yeast strains containing bait plasmid or appropriate controls. For clarity, descriptions of each plasmid are presented in Table E.

Table D: Clones isolated during yeast two-hybrid screen. Four independent clones were isolated for Troponin I, all others represent a single clone.

CLONE	DESCRIPTION
17.40 Decorin B	2 splice variants (A & B); widely expressed, found in connective tissues (ECM). Postulated to play a role in epithelial/mesenchymal transitions during organ development and shaping. (Danielson et al., 1993; Scholzen et al., 1994).
19.15 SPARC / Osteonectin / BM-40	Secreted Protein, Acidic, Rich in Cysteine; thought to be involved in cell proliferation, wound healing, modulation of ECM structure; so-called 'de-adhesion molecule'(Goldblum et al., 1994; Young et al., 1998)
16.27 Homology: COP9 Subunit 5/JAB1/MOV34 homologue (contains PCI domain)	Subunit of the multifunctional 26S proteosome; intracellular; clear physiological roles undefined (RNA binding, macromolecular assembly)(Asano et al., 1997; Hofmann and Bucher, 1998)
8.14 Cardiac Troponin I	Subunit of troponin complex, regulates contractility of striated muscle (cardiac; skeletal) (Bhavsar et al., 1996)
C1 "DAZ" associated protein	Deleted in Azoospermia; testes specific expression (Reijo et al., 1995)
17.26 Huntingtin-associated protein interacting protein (HAPIP)	Brain specific expression; contains tandem spectrin-like repeats (Colomer et al., 1997)
14.31 Ferritin light polypeptide	Light chain of ferritin; iron storage and detoxification complex (Harrison and Arosio, 1996)
15.22 SMAD/MADD homology [SMAD-4A]	Cytoplasmic/nuclear protein regulates transcription via low affinity DNA binding; specificity conferred by forming DNA-binding multimers (Wrana, 2000)
C2 Human Sarcoglycan; 50kD dystrophin associated glycoprotein / 'Adhalin'	Transmembrane glycoprotein; expression limited to skeletal, cardiac and some smooth muscle types (Roberds et al., 1993)

Table E: Yeast two-hybrid plasmid descriptions

PLASMID	DESCRIPTION
pGBKT7.Y2H2	Yeast two-hybrid bait construct used in library screen (see Figure 12)
pGBKT7	Empty yeast two-hybrid vector containing GAL4 DNA-Binding Domain. Constructs derived from pGBKT7 carry TRP1 allowing for growth on media lacking tryptophan.
pGADT7	Empty yeast two-hybrid vector containing GAL4 DNA-Activation Domain. Carries LEU2 for growth on media lacking leucine.
pGBKT7.Lam	Encodes a fusion of the Lamin C protein (known not to form complexes and shows rare and very limited interaction with other proteins) with the DNA-BD.
19.15	Library clone shown to contain cDNA encoding amino acids 119-272 (of 280) of SPARC
17.40	Library clone encoding amino acids 218 to 359 (of 359) of proteoglycan core protein Decorin

Initially, all mated yeast were grown on -Leu/-Trp plates to confirm the presence of both test plasmids in the diploid. The haploid strains Y187 and AH109 are auxotrophic for leucine and tryptophan. Growth on media lacking leucine and tryptophan indicates the presence of both plasmids. Combinations of plasmids found in diploid yeast (after mating) and growth profiles on –Leu/-Trp plates are summarized in Table F:

Table F: Growth of yeast containing yeast-two hybrid constructs and library clones on –Leu/-Trp media after mating

	pGBKT7.Y2H2	pGBKT7	pGBKT7.Lam
19.15	(+)	(+)	(+)
17.40	(+)	(+)	(+)
pGADT7	(+)	(+)	not done

In order to test whether any of the above combinations were sufficient to bring together both the DNA-AD and DNA-BD and drive expression of reporter genes, mated yeast were also plated on QDO (SD/-Ade/-His/-Leu/-Trp). As expected, only pGBKT7.Y2H2 and putative positive binding partners were shown to be sufficient for growth. The results of these tests are summarized in Table G.

As a further test of the potential binding partners identified during the library screen, mating between strain AH109 carrying 19.15 or 17.40 were performed with strain Y187 containing pGBKT7.Y2H2, pGBKT7 alone or pGBKT7.Lam, followed by growth on QDO or –Leu/-Trp. Colonies were then transferred to filters and *βgal* assays were performed. As expected, in the case of pGBKT7 and pGBKT7.Lam, no colonies were

Table G: Growth of yeast containing yeast-two hybrid constructs and library clones on QDO media after mating

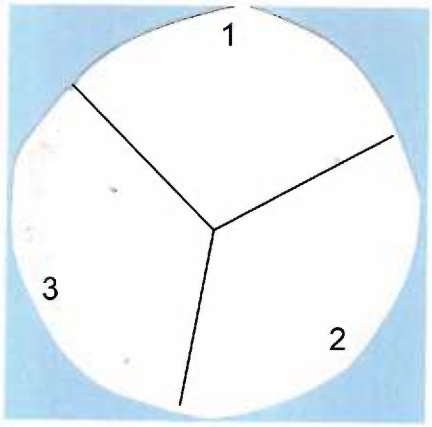
	pGBKT7.Y2H2	pGBKT7	pGBKT7.Lam
19.15	(+)	(-)	(-)
17.40	(+)	(-)	(-)
pGADT7	(-)	(-)	not done

obtained from QDO plates thus these *βgal* assays were carried out on filter lifts obtained from –Leu/-Trp plates. The results of these experiments indicate that the SPARC or Decorin DNA-AD fusion proteins interact with the CRELD1 DNA-BD fusion only. This outcome can be visualized in Figure 13.

Figure 13

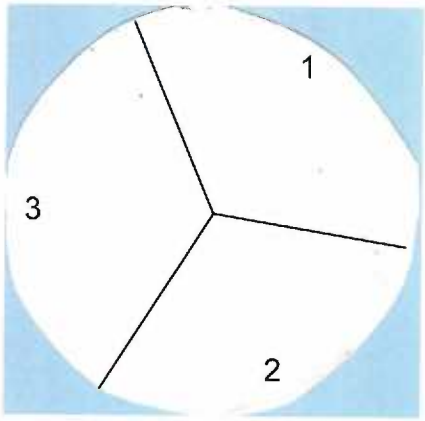
(A) Three individual colonies each containing both pGADT7 and pGBKT7.Y2H2 (bait construct), streaked on –Leu/-Trp plates followed by βgal assay, demonstrating the bait construct does not activate the reporter in the presence of empty AD vector. (B) βgal assay results from yeast containing both control vectors (empty DNA-BD and DNA-AD). (C) Library clone 19.15 (SPARC) in strain AH109 mated with strain Y187 containing either pGBKT7.Lam or pGBKT7 alone, followed by βgal assay (note: (-) is a section lacking yeast) (D) Clone 19.15 (SPARC) mated with pGBKT7.Y2H2 demonstrating robust color change after βgal assay, indicating interaction between CRELD1 and SPARC. (E) Results of assay with control strains mated with 17.40 (Decorin). (F) βgal assay using library clone 17.40 (Decorin) in strain AH109 mated with strain Y187 containing pGBKT7.Y2H2, again demonstrating interaction between bait and library inserts.

A



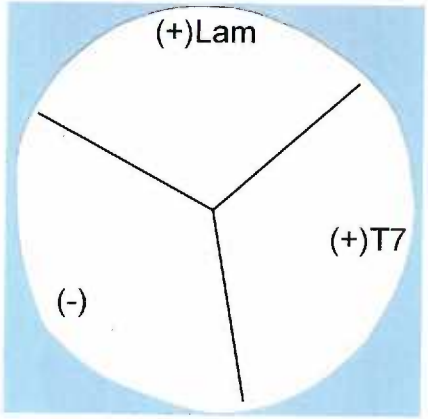
pGADT7 / pGBKT7.Y2H2

B



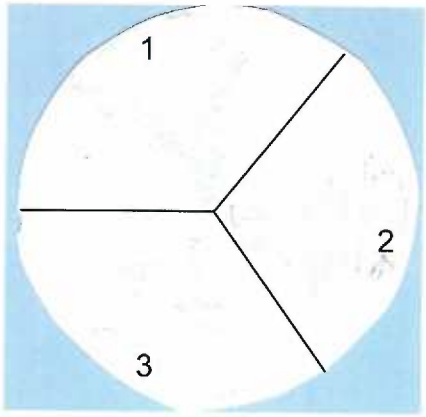
pGADT7 / pGBKT7

C



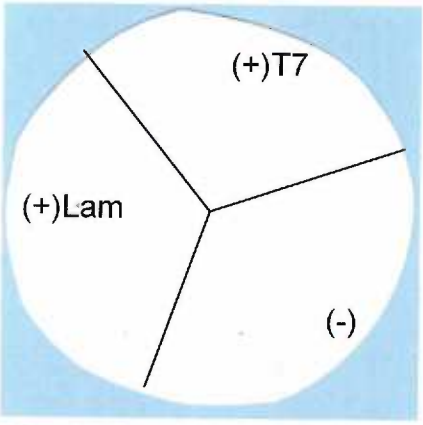
19.15 (SPARC)/ pGBKT7.Lam or pGBKT7

D



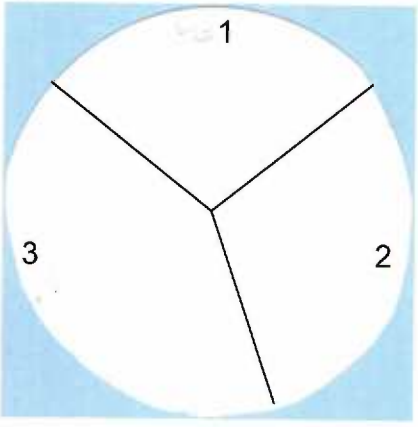
19.15 (SPARC) / pGBKT7.Y2H2

E



17.40 (Decorin) / pGBKT7.Lam or pGBKT7

F



17.40 (Decorin) / pGBKT7.Y2H2

***In vitro* transcription/translation and co-immunoprecipitation.**

In order to obtain labeled protein products for further analysis, an *in vitro* transcription/translation approach was taken. The vector pGBKT7, into which the CRELD1 fragment was cloned, contains a c-Myc tag 5' of the MCS. Similarly, library inserts were positioned 3' of an HA-tag. Sequences encoding both the c-Myc and HA epitopes were themselves 3' of a T7 promoter, allowing direct use of circular plasmid template with TNT T7 Coupled Wheat Germ Extract System reagents. We were able to successfully produce appropriately sized protein fragments for CRELD1, SPARC and Decorin labeled with ^{35}S-cysteine (data not shown).

To further substantiate interactions between CRELD1 and its potential binding partners, a co-immunoprecipitation approach was used. As described in materials and methods, bait and target translation products were mixed and incubated with either α-c-Myc or α-HA antibodies followed by addition of Protein A beads to precipitate complexes containing both protein fragments. SV-40 large T-antigen and p53 vectors were supplied with the kit for use as positive controls which worked as predicted (data not shown). Further, initial immunoprecipitation data suggested that the WE domain of CRELD1 formed complexes with SPARC and to a lesser degree Decorin. Several experiments were conducted where the ability to immunoprecipitate c-Myc tagged protein with α-c-Myc antibody verified the specificity of the interaction. The same approach was used for HA-tagged translated product. However, when α-HA antibody was used with c-Myc tagged protein (as a negative control), a significant amount of translated product was precipitated. This was also found to be the case with HA tagged products mixed with c-Myc antibody. Thus, while the results of these analyses were

suggestive, the high degree of background in negative control experiments prevent definitive conclusions from being drawn (data not shown).

CRELD1 GST-Fusion

Because the co-immunoprecipitation approach proved problematic, a secondary approach was chosen. This involved fusing the WE domain of CRELD1 to glutathione-S-transferase using the pGEX-2T vector, followed by expression in *E. coli*. After purification of fusion protein using GST resin, pull-down experiments were conceived where the CRELD1 fusion protein would be incubated with purified SPARC or Decorin, complexes isolated with a second step of purification over GST resin, followed by Western analysis with appropriate antibodies.

After transformation, induction and protein purification it was discovered that, while it was possible to detect GST-CRELD1 fusion protein by Western blotting, the quantities eluted off the GST resin were extremely small (data not shown). Furthermore, it was observed that the bulk of expressed protein formed insoluble inclusion bodies in *E. coli*. While numerous attempts were made to solubilize these aggregates and refold the expressed protein using 8M urea followed by stepwise removal of salts via dialysis, the latter stages of purification uniformly yielded an insoluble precipitate. Subsequent analysis revealed this precipitate was composed largely of expressed fusion protein. The viscous nature of this material, and the presence of multiple bands on Western blot lead us to postulate that the fusion protein was forming extensive disulfide bridges as urea was removed and this not only caused significant amounts of GST-CRELD1 to end up as an insoluble precipitate, but could also account for the very low yields of GST-CRELD 1

obtained from purification procedures using GST-resin and elution with excess glutathione. Due to these difficulties, insufficient quantities of properly folded and purified GST fusion protein were obtained, thus precluding use of this reagent in pull-down experiments. Nevertheless, denatured protein was detectable on Western analysis using α-GST antibodies, and we realized that it was possible to use this material to more completely characterize peptide antibodies previously generated in rabbit against CRELD1 (Rupp, Fouad et al., 2002). The use of this GST fusion protein for this purpose has been successful, and these antibodies are being used to identify cellular localization and tissue distribution of Creld1 at high resolution using immunohistochemical techniques.

Chapter IV

Discussion

Discussion

CRELD1 is a newly described gene encoding a highly conserved protein belonging to the EGF superfamily that maps to chromosome 3p25 (Rupp, Fouad et al., 2002). 3p- syndrome is a contiguous gene deletion syndrome characterized by loss of distal chromosome 3p, or interstitial deletion at 3p25-26. A critical region between markers D3S1263 (proximal) and D3S3594 (distal) has been defined. 3p- syndrome patients with more proximal breakpoints exhibit AVSD, whereas patients with more distal deletions do not (Mowery, 1993; Phipps et al., 1994; Drumheller et al., 1996). *CRELD1* expression is detected in the myocardium and endocardial cushions by *in situ* hybridization in chick embryos. The endocardial cushions are populated with mesenchymal cells that are derived specifically from the endocardium via an inductive signal generated by the myocardium. Mesenchyme proliferation and migration is essential for proper formation of the valves and septa of the mature four-chambered heart, and defects in these processes as well as in signaling and/or de-adhesion are predicted to result in AVSD.

Initially, FISH analysis demonstrated deletion of a single copy of *CRELD1* in 3p- syndrome patients with an associated cardiac malformation (Rupp, Fouad et al., 2002), and it has been shown by several different groups that haploinsufficiency of a single gene product can result in congenital heart defects (Ewart et al., 1993; Li et al., 1997b; Oda et al., 1997b; Schott et al., 1998). Subsequent examination of cell lines derived from 3p- patients without CHD using FISH revealed a *CRELD1* deletion in this subset of individuals as well (Rupp, Fouad et al., 2002). Based on this finding it is reasonable to

initially conclude *CRELD1* deletion is inconsequential with respect to the presence or absence of CHD. However, it is important to stress that incomplete penetrance has been documented in families with AVSD (Eldadah et al., 2001; Goldmuntz et al., 2001), thus it is possible that individuals with 3p- syndrome deleted for *CRELD1* do not universally exhibit cardiac anomalies. Taken together, these data led us to hypothesize that *CRELD1* may play a role in normal heart development and that *CRELD1* haploinsufficiency may be partly responsible for the heart abnormalities (AVSD) seen in 3p- syndrome patients.

In order to test this hypothesis, we initiated a project designed to recapitulate AVSD in mice by creating a mutation that results in a single *Creld1* null-allele. Further, we sought to understand on a molecular level the role of CRELD1. Thus, to gain additional insight into CRELD1 function, we performed a yeast two-hybrid interaction screen. To accomplish this, a CRELD1-GAL4 AD fusion protein was used in the hope of identifying interacting proteins for which functional data is available. This work details the successful generation of a *Creld1* knock-out mouse, and the identification of two small, multifunctional extracellular proteins implicated in mediating cell adhesion and migration as CRELD1 interaction partners.

Our laboratory is actively pursuing projects in parallel with those described here with the aim of demonstrating a clear role for *CRELD1* during cardiogenesis. One approach that has been particularly productive has been to take advantage of the Oregon Heart Registry, a catalogue of patient information that tracks individuals in whom heart malformation has been recognized. The availability of this resource has facilitated the collection of samples for *CRELD1* characterization. In addition, collaborations with

other laboratories have proved a valuable resource for the procurement of patient samples.

Mutation analysis has been performed using DNA from these sources. Of 63 total patients analyzed, 3 presented with syndromic CHD. An additional 12 demonstrated heterotaxy and CHD, while the remainder had isolated AV canal defects or AV canal defects with additional heart malformation. Oregon Heart Registry Patient 25778996, who has a partial AV canal, was found to have a heterozygous C to T transition at nucleotide 4201 resulting in the substitution of an arginine to cysteine. This mutation results in an additional cysteine in the second calcium binding EGF domain (Sue Robinson, Personal Communication). EGF domains contain specific disulfide bridges between cysteine residues, and the presence of an additional cysteine would be predicted to cause an alteration in disulfide bond distribution within the domain. That the structure of the molecule is affected by such a substitution has been demonstrated by altered mobility on native protein gel analysis following transient expression in 293 cells. Furthermore, it has been shown that this mobility shift is not due to either glycosylation or beta-hydroxylation (Sue Robinson, Personal Communication). Allele specific PCR has failed to detect this alteration in over 400 chromosomes. Similarly, Oregon Heart Registry Patient 15630221, an individual with a partial AV canal exhibits a heterozygous G to A transition at nucleotide 37, which results in methionine replacing valine at amino acid position 13. Computer analysis predicts such a substitution abolishes the secretion signal of CRELD1. Normal chromosomes are currently being screened for this alteration.

Samples received from collaborators have been of equally informative. Patient G9 who has a complete AV canal defect was found to harbor a heterozygous C to T transition at nucleotide 4148. Not unlike Oregon Heart Registry patient 25778996, this mutation occurs in the second calcium-binding EGF domain, though rather than introducing cysteine, substitutes an isoleucine residue for threonine at amino acid 311. Again, this alteration was not found in 382 normal chromosomes by allele specific PCR. Most interestingly, patient G7, who presents with an AV canal and heterotaxy shows a heterozygous G to A transition at nucleotide 1566 which predicts the substitution of arginine with histidine at amino acid 107. This substitution is positioned within the unique WE domain of CRELD1. Whether or not this has implications for the clinical finding of heterotaxy in this individual remains to be determined. Further, this individual is of mixed African-American (Caribbean) and Hispanic ethnicity, and an analysis of over 100 chromosomes from each population has thus far failed to demonstrate that this substitution is a polymorphism. The discovery that *CRELD1* mutations occur in individuals with CHD, but not in many hundreds of appropriately selected normal control samples is perhaps the strongest evidence yet that C*RELD1* plays a critical role in cardiogenesis. Recapitulation of CHD in a mouse model should, in conjunction with the human mutation data, show unequivocally that CRELD1 is an essential for heart development.

CRELD1 interacts with Decorin and SPARC in a yeast two-hybrid system

Two CRELD1 binding partners have been identified by conducting a yeast-two hybrid screen using a portion of the human CRELD1 protein fused to a GAL4 activation domain. Sequence analysis has revealed that CRELD1 is a modular protein, containing a stereotypical secretion signal, a short, low-complexity proline rich domain, tandem EGF-domains, two calcium binding EGF-like domains and is predicted to be anchored to the cell membrane by a pair of transmembrane spanning helices. The protein also contains a tryptophan and glutamic acid rich region (WE domain), which lies at the amino-terminal end of the molecule, just downstream of the proline rich tract.

When homology searches were conducted using CRELD1 sequence, no significant hits were identified using the WE domain in unrelated molecules. However, during these searches, orthologous and homologous genes were identified from several species. Comparisons between mouse, human, Chinese hamster, *C. elegans* and *D. melanogaster* showed that, though there is no significant similarity of the WE domain between CRELD1 and other proteins, the WE domain is very highly conserved between species (Appendix II). This suggests a strong evolutionary pressure to maintain intact function. Because conservation across the WE domain is high, and because it is apparently unique to CRELD proteins, this region was chosen as 'bait' for a yeast two-hybrid library screen.

This screen resulted in the identification of two putative interacting proteins, decorin and SPARC (secreted protein, acidic, rich in cysteine). SPARC and decorin are well-studied, multi-functional extracellular (or matricellular) molecules. Though other

interacting proteins were identified, they were excluded from further study based on two principle criteria: sub-cellular localization (nuclear or intracellular rather than extracellular) and limited, tissue-specific expression patterns.

Decorin

Decorin is a proteoglycan first isolated from adult bovine cartilage in 1985 (Rosenberg et al., 1985). Decorin, which consists of a ~ 36 kD protein core, is a major constituent proteoglycan of most connective tissue and belongs to the family of small leucine-rich proteoglycans (SLRPs) that are secreted into the ECM (Scholzen et al., 1994; Iozzo, 1999). The SLRP family currently includes nine proteins. Decorin and biglycan are the most closely related SLRPs and contain a propeptide that may serve as a recognition signal for enzymes involved in the addition of glycosaminoglycan chains. Other SLRP family members include fibromodulin, lumican, epiphican, keratocan, PRELP, PG-Lb and osteoglycin (Iozzo, 1999) The hallmark feature of decorin and biglycan is a canonical 24 amino acid residue leucine-rich region with asparagine and leucine positionaly conserved ($LX_2LXLX_2NX(L/I)$) and repeated within the protein multiple times (Scholzen et al., 1994; Iozzo, 1999). Decorin is known to bind a diverse group of molecules including TGFβ, several fibrillar collagens, fibronectin, thrombospondin as well as cell endocytosis receptors (Yamaguchi et al., 1990; Schmidt et al., 1991) (Winnemoller et al., 1991) (Winnemoller et al., 1992) (Hausser et al., 1998; Iozzo, 1999). Additionally, decorin has been shown to bind tropoelastin (the precursor to elastin) as well as fibrillin-containing microfibrils, suggesting decorin may play a role in

elastinogenesis (Reinboth et al., 2002). Decorin acts to attenuate TGFβ activity via direct binding of the growth factor (Yamaguchi et al., 1990). Moreover, decorin has been shown to exert cytostatic effects in a variety of tumor cell lines when stably transected with decorin (Santra et al., 1997). More recently, decorin has been found to induce cell cycle arrest via direct interaction with EGFR, leading to phosphorylation of MAP kinase and p21 induction in A431 squamous carcinoma cells (Iozzo et al., 1999b). Decorin has also been shown to up-regulate p21 in endothelial and epithelial cells, though through a separate pathway (Schonherr et al., 2001).

Decorin has been cloned from several species and comparison of amino acid sequence demonstrates a high degree of conservation (>80%) (Scholzen et al., 1994). Expression patterns of decorin in the adult mouse have been probed by Northern blot. Decorin shows widespread expression in mature animals, with strongest levels in skin, liver and in heart, specifically the pericardium. This is in contrast to the signals detected in developing embryos using *in situ* hybridization, where a dynamically changing pattern is seen. During development, decorin expression is initially found in neural cell precursors (e11), later becoming restricted to the linings of the major organs, including pericardium of the heart and meninges (e13). Signals were subsequently seen in connective tissues and skin, as well as the gut (e16). The common theme underlying the observed shifts in expression is that decorin expression occurs in the embryonic mesenchymal tissues and organ-lining layers. Based on these expression patterns, the authors speculate decorin may play a role in epithelial-mesenchymal interactions during organogenesis (Scholzen et al., 1994).

In order to gain further insight into the functional role of decorin, mice harboring a disruption at the decorin locus have been generated (Danielson et al., 1997). Mendelian ratios were seen in offspring of decorin deficient heterozygotes, demonstrating decorin is not required for survival. Though gross anatomical inspection was normal, a severe fragility of skin, marked by dermal thinning and loose connective tissue in the hypodermal layer was found. Biomechanical analysis of skin segments showed a three-fold reduction in tensile strength compared to wild-type littermates. Interestingly, this phenotype was not fully penetrant. Ultrastructural analysis of dermal collagen by electron microscopy revealed a disorganized collagen fibril array. Further, collagen fibrils showed irregularity in shape and size, both longitudinally as well as in cross-section. No abnormalities of internal organs were reported (Danielson et al., 1997).

Because decorin has been shown to induce cell growth arrest, but decorin-deficient mice show no sign of unchecked cell proliferation, Iozzo and colleagues hypothesized that the absence of decorin could lead to tumorigenesis in a suitable genetic background. This hypothesis was tested by making a mouse mutant for both decorin and the well-known tumor-suppressor p53 (Iozzo et al., 1999a). Double homozygous mutant mice were uniformly found to develop thymic lymphomas, and tumorigenesis occurred more rapidly than in the p53 knockout alone. A p53 $^{-/-}$ decorin $^{-/+}$ mutant showed thymic lymphomas with the same frequency, though at a much slower rate. Growth of cultured lymphoma cells was slowed considerably with the addition of exogenous decorin or decorin secreted from fibroblasts, and the human decorin protein core was shown to be the functional equivalent of mouse protein. In fact, isolated human protein core was shown to be a more potent inhibitor of proliferation, suggesting that this activity is

mediated by the protein itself and not the glycosaminoglycan side chain. In a separate study, exogenous addition or ectopic expression of the isolated protein core was found to cause growth suppression in neoplastic cells of various origins (Santra et al., 1997). It was further found that growth arrest in these cells was p21 dependent, and de novo expression of decorin in HCT116 human colon carcinoma cells with a disrupted p21 gene were unresponsive. The conclusion drawn from these data is that decorin can act to negatively regulate cell proliferation.

Using an *in vitro* assay designed to test the response of endothelial cells to different ECM substrates, it was found that decorin alone (isolated from cartilage), or decorin with thrombospondin-1 was able to inhibit the formation of endothelial tube-like structures (TLS) (Davies et al., 2001). TLS formation is a measure of neovascularization activity (Davis and Camarillo, 1996). Further, decorin was able to prevent the formation of endothelial cell 'aggregates' when added to cell suspensions during or prior to plating. This suggest that decorin plays an 'anti-adhesive' role via interactions involving anchoring extracellular substrate(s) and/or cell surface receptors.

When bovine endothelial cells were transduced with replication-deficient retroviral vectors containing bovine decorin, it was found that these cells show decreased cellular migration during monolayer outgrowth assays when compared to cells transduced with empty vector alone (Kinsella et al., 2000). The edges of these monolayers did appear more dispersed after 72 hours, again suggesting decorin has de-adhesive properties. Accelerated fibrillogenesis was a consistent feature of the expressing cells, and this was found to be dose-dependent when non-decorin expressing cells were incubated with increasing amounts of purified protein. These investigators also studied

the effect of decorin on cell proliferation by BrdU labeling, and found that endothelial cell proliferation is unaffected by decorin, which seemingly contradicts the finding that decorin causes cell cycle arrest (Iozzo et al., 1999b). Clearly, more work needs to be done to clarify these conflicting results, and it is likely that the functional outcome of decorin activity is modulated by the extracellular matrix and cell surface microenvironment.

SPARC

SPARC (also known as osteonectin or BM-40) encodes a highly conserved developmentally regulated matricellular protein belonging to a group of factors that are thought to mediate cell-matrix interactions rather than contributing directly to ECM structure (Murphy-Ullrich, 2001). Other members of this class include thrombospondins 1 and 2, tenascins C and X, and osteopontin (Brekken and Sage, 2001). SPARC is a 32 kD (286 amino acid) glycoprotein is composed of three 'modules' including an N-terminal acidic domain, a cysteine rich 'follistatin-like' domain (FS) and an extracellular calcium-binding domain (EC) (Hohenester et al., 1997). Like decorin, SPARC is multifunctional and is reported to mediate embryogenesis, tissue remodeling and wound healing, modulate cell shape and adhesion as well as act to enhance or inhibit cell proliferation (Bradshaw and Sage, 2001) (Brekken and Sage, 2001; Murphy-Ullrich, 2001).

SPARC was originally purified from developing bone and was thus thought to be bone specific, as it was shown to bind calcium with high affinity, perhaps acting as a link

between mineral and collagen (Termine et al., 1981). Subsequent observation showed that SPARC interacts with a wide variety of extracellular molecules including platelet derived growth factor (PDGF), vascular endothelial growth factor (VEGF), and TGFβ. Furthermore, SPARC is reported to interact with structural matrix molecules such as collagens type I, III, IV and V as well as vitronectin and entactin/nidogen, again highlighting the varied roles of this protein (Bradshaw and Sage, 2001) (Brekken and Sage, 2001).

The modular nature of SPARC has lead to dissection of the functional and structural domains of the molecule. The crystal structure for the globular EC domain of SPARC has been solved at 2.0 Angstrom resolution (Hohenester et al., 1996). The solution revealed a tandem pair of canonical EF-hand calcium-binding sites, as well as three α-helices postulated to constitute the collagen-binding site. In a subsequent study, the crystal structure of FS domain of SPARC was elucidated (Hohenester et al., 1997). The FS domain is stabilized by five disulfide bonds and takes a more open conformation than that of the EC domain. Sequence alignment with other follistatin-like proteins (follistatins I, II and III; agrins) shows nearly complete conservation of the five disulfide cysteine pairs. These domains are reminiscent of the 'Kazal' family of serine protease inhibitors, suggesting a possible relationship between these molecules, though this activity has never been attributed to SPARC. Structure definition has shown an inhibitory loop from the EC domain extends into the protease inhibitor-like FS domain, which may explain the failure to detect inhibition of proteases by SPARC. The N-terminal acidic domain is short (~50 amino acids) and presumably loosely ordered, thus

the FS-EC structure provides a near-complete picture of SPARC architecture (Hohenester et al., 1997).

An *in vitro* system of angiogenesis was used to study the response of endothelial cells to the addition of SPARC (Lane et al., 1992). Bovine aortic endothelial cell clones were isolated due to their ability to spontaneously form 'tubes' and 'cords' after ECM synthesis. Contact-inhibited endothelial cells, which do not form these structures, were used as controls. It was found that exogenous SPARC added to neovascularizing cells decreased the synthesis of both fibronectin and thrombospondin, but had no effect on collagen I. Peptides derived from SPARC synthesized to represent different portions of the protein were used to identify functional domains. A peptide from the acidic region of SPARC (amino acids 4-23) was found to substitute for the intact protein in this assay, although the peptide was considerably less bioactive than intact protein. Nevertheless, a 19 residue peptide containing an identical but scrambled amino acid profile as peptide 4-23 was entirely inactive, suggesting this region of SPARC does contain biological activity. Higher concentration of intact protein caused changes in cell morphology (rounding) in addition to shifting levels of extracellular molecule expression. These effects were not seen in confluent cultures free of 'tubes' and 'cords'. Though the authors speculate that these observations were likely the result of SPARC-induced changes in cell adhesion to the ECM or changes in the ECM itself, they could not rule out the existence of a cell surface SPARC receptor.

Subsequent experiments conducted by the same laboratory sought to detect direct interaction between SPARC and endothelial cells via a putative receptor (Yost and Sage, 1993). In order to carry out these experiments, ^{125}I-SPARC was added directly to

endothelial cells followed by washing and detection. It was found that SPARC can bind directly to endothelial cells. Predictably, SPARC adhesion to these cells is calcium dependent (1mM Ca^{++} at pH 7.1), though surprisingly stable through a relatively broad range of pH (pH 4 – 7). In this study peptide 4-23 was unable to compete with binding. Thus, other peptides were prepared and analyzed in competition experiments. In this manner it was found that a peptide containing amino acids 254-273 was able to compete directly with full length SPARC. This sequence localizes to the EF-hand region in the EC domain of the mature protein. Peptide 254-273 was immobilized to a solid support and used in affinity isolation experiments with proteins recovered from endothelial cell membrane preparations. By this method several proteins were purified ranging in size from 100 to 153 kD. However, the identities of these fragments remained unclear and the authors caution that in general, short synthetic peptides or protein fragments often show binding properties different from native protein.

In a follow-up study, Murphy-Ullrich and co-workers were able to show that SPARC protein purified from mouse or recombinant protein produced in *S. cerevisiae* caused loss of cellular focal adhesion plaques in endothelial cells (Murphy-Ullrich et al., 1995). This effect could be blocked by using rabbit anti-SPARC polyclonal antibody. Interestingly, though cells appeared round and lost focal adhesion plaques with the addition of SPARC, the integrin receptor and its ECM ligand (vitronectin) were shown to remain intact. It has been suggested that this represents an intermediate state of adhesion, one in which cell migration is favored (Murphy-Ullrich, 2001). As before, peptides derived from distinct regions of SPARC were analyzed. Two peptides, 54-73 and 254-273 were found to cause dissolution of focal adhesion plaques, though this activity was

not additive. Furthermore, as with the intact protein, this activity could be blocked with antibodies raised against these peptides. The authors highlight the fact that though each peptide can cause loss of focal adhesion plaques, they each contain properties of their own, and it is not clear whether these peptides are recognized by a single or multiple receptors.

In an unrelated set of experiments, it was found that heart-conditioned medium derived from newborn rat ventricular primary cell culture was able to cause focal adhesion disassembly in rat embryo fibroblasts (Dunlevy and Couchman, 1993). This work also showed that stimulation with known growth factors (PDGF, FGF, EGF, TGFβ) could not substitute for conditioned heart medium. The authors concluded that highly migratory fibroblasts from newborn heart secrete a factor(s) that mediates focal adhesion stability, and that the 'wounded' heart tissue from which the primary cultures were derived enhanced release of these factors. Though follow up work by Dunlevy and colleagues implicated the cytokine IL8 in focal adhesion stability in this experimental system, it is tempting to speculate that SPARC may be a secondary factor involved in this process, particularly given that IL8 was not as effective as complete conditioned heart medium (Dunlevy and Couchman, 1995). In fact, SPARC has been shown to become upregulated (>50%) in the ventricular myocardium of hypertrophic rat heart after infusion of isoproterenol (a model of myocardial remodeling after 'wounding')(Masson et al., 1998). Further, the increase in SPARC expression was accompanied by enhanced expression of known SPARC binding partners collagens I and III.

As with decorin, SPARC knock out mice have been generated (Gilmour et al., 1998; Norose et al., 1998). SPARC heterozygotes mutants appeared normal and fertile,

as did homozygous offspring, which were obtained in expected Mendelian ratios (n=95). Northern and western analysis confirmed that SPARC production was disrupted. Furthermore, though the authors report an established role for SPARC in morphogenesis of bone, muscle, blood vessels, skin and heart, these and other tissues were normal by histological exam. Bone structure as assessed by alizarin red staining was indistinguishable between SPARC $^{-/-}$ and SPARC$^{+/+}$ littermates on an MF1 background. These findings were unexpected given the widespread expression of SPARC during development.

The major finding in SPARC null animals was the unexpected development of bi-lateral cataracts at ~ 6 months of age (Gilmour et al., 1998). Histology revealed posterior rupture of the lens and abnormal epithelial cell migration and infiltration of this region. Closer exam by electron microscopy found disorganized lens fibers and unusual epithelial cell rounding and vacuolation when compared to epithelial cells from the normal mouse. The mechanism by which this occurs is unresolved. ECM components including collagen IV, laminin 1, perlecan, and entactin were examined by immunohistochemistry and no changes were identified between mutant and control animals.

Closer exam of SPARC null mouse skeletons showed that mutant animals did in fact show subtle changes in bone physiology when bred on a 129SV/C57BL/6 background (Norose et al., 1998; Delany et al., 2000). Vertebrae density from 11 week old mutants was found to be decreased via bone density scan. This density decrease was also found in the long bones (tibia and femur), and worsened with time. When the biomechanical properties of bones from these mice were examined, it was found that

bones of SPARC null mice were considerably weaker. These findings were attributed to low-turnover osteopenia, given that the number of osteoblasts and osteoclasts were decreased.

Primary mesenchymal cells have been isolated from SPARC$^{-/-}$ mice and studied *in vitro* in order to determine whether primary cells lacking SPARC looked and behaved differently than wild-type counterparts. Additionally, SPARC null cells were tested to see if they could respond to exogenous SPARC in the absence of endogenous production (Bradshaw et al., 1999). While morphological changes were not readily apparent in aortic smooth muscle cells or skin fibroblasts, mesangial cells showed an increase in focal adhesions and increased actin fibers located at the cell periphery. Interestingly, mesangial cell morphology did not change with the addition of exogenous SPARC, though skin fibroblasts showed rounding. Further, immunolocalization of SPARC showed it was distributed only at discrete sites on the cell surface. In this instance, normal and mutant cells behaved in an identical fashion. All three mutant cell types showed an accelerated rate of proliferation compared to controls. This difference was eliminated with the addition of SPARC to the cultures.

Because SPARC has been implicated in the regulation of cell adhesion and migration, it was hypothesized that the protein may be a modulator of tumor invasiveness. Evidence supporting this notion was demonstrated *in vitro* using SPARC antisense RNA and human melanoma cells (Ledda et al., 1997). Cell lines expressing the antisense message were found to have reduced capacity for adherence and invasion of matrigel membranes in a dose dependent manner. In addition, tumors with or without SPARC antisense RNA were xenografted to nude mice. Mice with tumors previously

transfected with antisense vectors remained tumor free after 8 months, whereas tumors were found in all mice injected with unaltered cells (Ledda et al., 1997).

This idea was examined in further detail by Rempel and others by comparing nonrecurring benign, noninvasive meningiomas with aggressive, invasive recurrent meningiomas (Rempel et al., 1999). Tumors were surgically removed from patients, graded and examined by immunohistochemistry for SPARC. Noninvasive histologicaly benign tumors were found to be SPARC negative, whereas all invasive tumors, whether benign or malignant, were found to be SPARC positive at the normal tissue-tumor interface. Thus, the authors suggest SPARC be considered as a candidate marker to identify potential or invasive meningiomas.

CRELD1 as a putative modulator of cell adhesion

As mentioned previously, examination of public online databases revealed *CRELD1* orthologous from a variety of organisms, including *C. elegans, D. melanogaster*, rodents, and primates. Commercially available multiple species Southern blots probed with *CRELD1*-specific probes have demonstrated *CRELD1* is present in mammals and birds, but not yeast (Rupp, 1999). Localization of *CRELD1* message by whole mount *in situ* hybridization and Northern blot has shown CRELD1 expression in multiple developing tissues, including heart, skeletal muscle, limb buds, branchial arches and brain (Rupp, Fouad et al., 2002). Higher resolution examination of heart tissue in chick embryos revealed expression in the endocardial cushions, endocardium and myocardium. Further, *CRELD1* is deleted on one chromosome 3 homolog in 3p-

syndrome patients with AVSD. Taken together, these data led us to hypothesize that *CRELD1* expression is essential for normal cardiogenesis.

Speculation with regard to possible mechanism(s) by which *CRELD1* haploinsufficiency could lead to AVSD drew attention to EMT in heart. Defects in endocardial cushion EMT are predicted to result in congenital AV canal and OT defects (Olson and Srivastava, 1996). Furthermore, components of the extracellular matrix have been shown to play a significant role in this process (Boyer et al., 2000; Kiemer et al., 2001). Sequence analysis of CRELD1 showed features characteristic of extracellular proteins, including an amino-terminal signal sequence as well as tandem EGF-like and calcium binding EGF-like domains. Additionally, there is evidence that alternative splicing at the 3' end of the message results in a protein lacking a transmembrane spanning domain predicting a soluble molecule secreted from cells (Rupp, Fouad et al., 2002). Thus, we previously proposed that CRELD1 could be playing a role in signaling, specifically as part of the soluble particulate matrix. An inductive signal is secreted by the myocardium and initiates delamination of endocardial cells to migrating mesenchyme that then proliferate in the cardiac jelly of the endocardial cushions. These cells ultimately differentiate, giving rise to valve and septal tissues (Eisenberg and Markwald, 1995). The EDTA-soluble inductive signal is a multi-protein complex known to contain fibronectin, ES130, hLAMP and transferrin as well as other unidentified components (Krug et al., 1987; Mjaatvedt and Markwald, 1989; Rezaee et al., 1993; Sinning, 1997).

Evidence that CRELD1 interacts with SPARC and decorin, two small ECM molecules known to mediate cell-cell and cell-matrix interactions, has led us to consider the possibility that CRELD1 acts to mediate cell adhesion rather than playing a direct role

in inductive signaling as a soluble ligand. However, this would not preclude a role for CRELD1 in EMT. There are several lines of evidence that support the idea CRELD1 is involved in cell adhesion. First, the presence of EGF-like domains in molecules at cell-cell and cell-substrate interfaces, including well known families such as the cadherins and laminins, as well as in more recently discovered adhesion molecules, is well established (Hynes and Zhao, 2000; Morimura et al., 2001). Second, homologous and orthologous *CRELD1* genes have been found in a variety of metazoans, though not in single-celled eukaryotes. Third, *CRELD1* expression is widespread; though most intense in heart and skeletal muscle, all tissues examined express varying levels of *CRELD1* message (adult brain, lung, liver, kidney, pancreas; placenta, fetal brain, lung, liver and kidney). This suggests a more global, and perhaps homeostatic role rather than a more spatially and temporally limited activity, as would be the case with inductive signaling during development. Fourth, a group of experiments in COS-7 cells where amino and carboxy-terminal FLAG – epitope tagged CRELD1 constructs were performed (data not shown). When an α-FLAG antibody was used for immunocytochemical localization in both fixed and unfixed cells, CRELD1 was predominantly found at both the cell periphery and within the endoplasmic reticulum and Golgi apparatus. While these procedures were not pursued to a definitive conclusion, the initial evidence supports the notion that at least some CRELD1 protein localizes to the membrane. Similarly, FLAG-epitope tagged protein can be detected in transient and stably transfected 293 cell membrane preparations by immunoblot providing direct evidence that CRELD1 is a membrane component (Sue Robinson, Personal Communication). Finally, through an extensive comparison of known cell adhesion proteins and common adhesion modules against the

euchromatic genomic sequences of *Drosophila melanogaster*, Hynes and Zhao independently identified the *CRELD1* ortholog as likely involved in cell adhesion (Supplemental table S6; CT# 31762 at: http://www.jcb.org/cgi/content/full/150/2/F89/DC1) (Hynes and Zhao, 2000).

As mentioned, though CRELD1 may not serve as an inductive signal, its expression pattern is provocative with regard to EMT. The distribution of *CRELD1* message detected by *in situ* hybridization in embryos shows most prominent levels in heart, limb buds, brain and branchial arches (Rupp, Fouad et al., 2002). This shares striking overlap with areas reported to undergo EMT inductive interactions during development, specifically Henson's node, somitic myotome, the branchial arch apparatus, under the limb ectoderm and neural crest (Spence et al., 1992; Sinning, 1997). Similarly, both SPARC and decorin are widely expressed during development, particularly in areas of tissue remodeling and later during wound healing (Scholzen et al., 1994; Bradshaw and Sage, 2001). Further, these molecules are both known to precipitate alterations in cell shape and confer a 'de-adhesive' phenotype to cells (Brekken and Sage, 2001; Davies et al., 2001). SPARC and decorin are also known to interact with vitronectin and fibronectin respectively, both of which are crucial factors in EMT progression (Schmidt et al., 1991; Bouchey et al., 1996; Rosenblatt et al., 1997).

The sequence of events leading up to EMT and differentiation are only partly understood. There is currently no direct evidence that SPARC or decorin are components of the particulate matrix, nor evidence demonstrating they are involved in EMT. However, based on their patterns of expression, their de-adhesive properties, and their soluble nature, both of these proteins are reasonable candidates for involvement in this

process. SPARC and decorin both exhibit de-adhesive properties which would suggest a primary role in epithelial disassembly – a hallmark of early EMT. However, these molecules are also implicated in slowing proliferation and promoting differentiation - events which typically define the end of EMT. In fact, SPARC expression is reported to decrease dramatically in differentiated tissues, becoming limited mainly to tissues undergoing constant turnover (Bradshaw and Sage, 2001; Brekken and Sage, 2001). Interestingly, this parallels the behavior of the particulate matrix, which can only be detected during the initial steps of EMT and disappears as cells begin to differentiate (Mjaatvedt et al., 1991). It is worth stressing that the functional consequences of localized SPARC and decorin expression are almost certainly dependent on cellular integration of complex (perhaps even conflicting) signals from their immediate environment (Giancotti and Ruoslahti, 1999; Iozzo, 1999). How these signals are sorted is an area of investigation about which much has yet to be learned. Future work will determine if SPARC or decorin are involved in EMT.

A cell surface receptor(s) for SPARC has been postulated, though to date no such receptor(s) has been identified (Yost and Sage, 1993; Bradshaw and Sage, 2001; Murphy-Ullrich, 2001). Furthermore, structural studies have proposed the putative-receptor binding epitope of SPARC is found primarily within the EC domain (Hohenester et al., 1997). During the yeast two-hybrid screen, the WE domain of CRELD1 was shown to interact with a SPARC clone encoding amino acids 119 to 272 (Figure 14).

Figure 14

Structure of SPARC protein. A ribbon diagram derived from crystallographic data shows the three modular domains of SPARC. Representative activities attributed to each domain are shown beneath the designated amino acids. A yeast two-hybrid clone (19.15) was isolated encoding amino acids 119-272 of SPARC. This fragment includes nearly all of the EC domain, which has been shown to bind to cells.

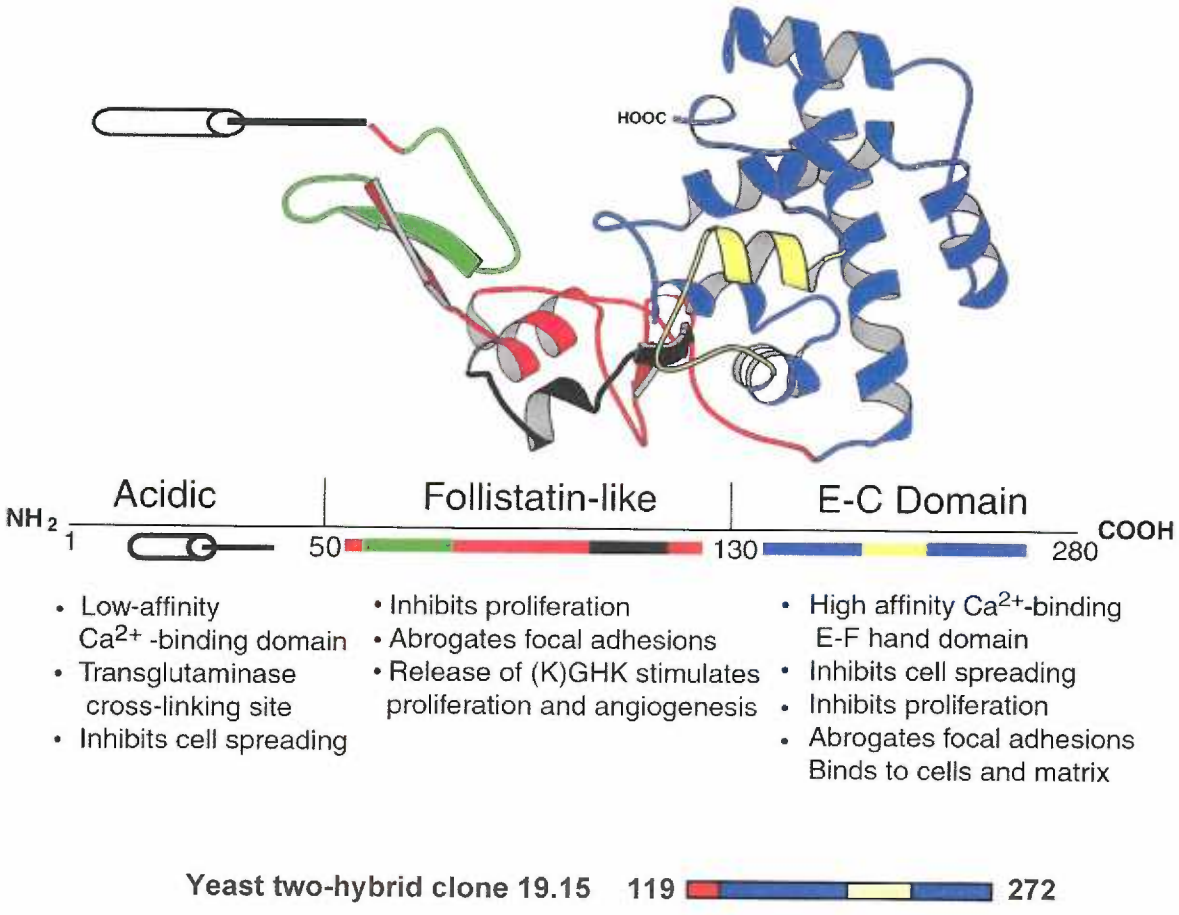

HOOC
Acidic
Follistatin-like
E-C Domain
NH2
1
50
130
280
COOH
• Low-affinity Ca2+ -binding domain
• Transglutaminase cross-linking site
• Inhibits cell spreading
• Inhibits proliferation
• Abrogates focal adhesions
• Release of (K)GHK stimulates proliferation and angiogenesis
• High affinity Ca2+-binding E-F hand domain
• Inhibits cell spreading
• Inhibits proliferation
• Abrogates focal adhesions
Binds to cells and matrix
Yeast two-hybrid clone 19.15
119
272

Figure 14

This region of the protein nearly encompasses the complete SPARC EC domain, and includes all but one amino acid of peptide 254-273 which has been shown to compete with full-length SPARC for binding to endothelial cells (Yost and Sage, 1993). Binding of SPARC to the cell surface is thought to be calcium dependent and occur through the calcium-binding EF-hand region found within the EC domain (Hohenester et al., 1997). Given the apparent requirement of calcium for CRELD1 structure (in stabilizing tandem calcium-binding EGF-like domains) and localization to the cell membrane, one attractive possibility is that CRELD1 is a SPARC receptor.

An alternative option is that SPARC and decorin act with CRELD1 in such a way that putative CRELD1 cell adhesive activity is reduced, and proteins interacting with CRELD1 supporting cell adhesion remain undiscovered. It is important to stress both SPARC and decorin share several common binding partners (i.e. collagen types II, III, IV, TGFβ, thrombospondin 1) though there are no reports of structural similarity between the two. Thus, it seems plausible that SPARC and decorin interaction with CRELD1 are not mutually exclusive. Regardless, at this stage such scenarios are purely speculative and more work needs to be done.

Several approaches were used in a further attempt to demonstrate interaction between the CRELD1 WE domain and SPARC or decorin. Primarily, these experiments were dependent on the incorporation of [^{35}S]-cysteine into proteins produced by *in vitro* transcription and translation (TNT) and autoradiography following interaction assays. Co-immunoprecipitation experiments with C-myc epitope tagged CRELD1 WE domain and HA epitope tagged SPARC, or HA-decorin did not differ significantly from controls (data not shown). Further, though GST-WE domain was successfully expressed in *E. coli*

(detectable on immunoblot), attempts to "pull down" SPARC or decorin were unsuccessful (data not shown). One possible explanation is that this reflects alteration in tertiary structure of the GST-WE domain fusion. This is a reasonable assumption given the observation that overexpression of this fusion protein in bacteria resulted in the formation of inclusion bodies, and protein recovery was dependent on extraction using 8M urea followed by dialysis in the presence of DTT. Furthermore, dialyzed GST fusion protein was found to adhere to glutathione resin with relatively low affinity. Thus, demonstration of SPARC and decorin interaction with CRELD1 will require the use of other, more sensitive assays where protein structural integrity is more likely to be maintained, such as in transfected cells.

Creld1 knockout mouse colony establishment

We hypothesized that haploinsufficiency of *Creld1* in mice would result in congenital heart malformation based on the finding that *CRELD1* is haploinsufficient in patients with 3p- syndrome and an associated heart defect. In order to test this hypothesis we proposed making heterozygous *Creld1* mutant mice followed by assessment for CHD. This work details the successful cloning, characterization and targeted deletion of murine *Creld1*. Chimeras have been generated using ES cell lines containing a *Creld1* null-allele. This allele is transmitted through the paternal germ line. Expansion of this mouse colony is now underway, and our initial data shows that *Creld1* heterozygotes are normal at birth without evidence of a heart annomoly. Interestingly, homozygous *Creld1* mice expire *in utero*. The etiology of this finding is unclear. Because this project is based

largely on the finding that *Creld1* is deleted in 3p- syndrome, (a contiguous gene deletion syndrome) useful parallels can be drawn from studies among the most common contiguous gene deletion disorders, the velocardiofacial (OMIM#192430) and DiGeorge syndromes (OMIM#188400).

Often referred to collectively as *del22q11* syndrome, DiGeorge syndrome (DGS) and velocardiofacial syndrome (VCFS) are so-called 'genomic disorders' marked by a recurring DNA rearrangement that is due in part to local genomic architecture predisposing the region to alteration (Stankiewicz and Lupski, 2002). The *del22q11* syndrome is defined by a typical 3 Mb deletion in ~ 90% of cases and exhibits both incomplete penetrance and variable expressivity. However, much smaller deletions and balanced translocations result in phenotypes remarkably similar to those in individuals harboring more extensive deletions. These include congenital cardiovascular defects (Tetralogy of Fallot; OT defects), hypocalcemia, thyroid and parathyroid aplasia/hypoplasia, characteristic facies (wide-set eyes, broad nasal root, small jaw, cleft palate), ear abnormalities, as well as behavioral (learning and memory deficit) and sometimes psychiatric disorders (Lindsay, 2001). It widely thought that haploinsufficiency of one or more genes from the region gives rise to the phenotype, though identification of the specific genes involved in the pathogenesis of *del22q11* syndrome has proved problematic, since patients with non-overlapping deletions show such similar anomalies (Lindsay, 2001). In more recent years, the use of mouse models to dissect the underlying genetic defects and their phenotypic consequences has been of particular utility.

A region of synteny between 22q11 in humans and mouse chromosome 16 has been characterized (Sutherland et al., 1998). Though gene order is not conserved across this region, at least 19 murine orthologous genes from human chromosome 22q11 were identified. Using this information, a mouse model of *del22q11* was made using Cre-*loxP* chromosome engineering to delete 18 of 19 homologs in mice (referred to as *Df1* mice) (Lindsay et al., 1999). Eighteen percent of adult mice with the *Df1/+* genotype had CHD. Examination of embryos at e11.5 showed nearly half harbored a heart defect indicating *in utero* lethality. Crossing *Df1/+* mice with mice engineered to have a duplication of this syntenic region (thus restoring gene dosage) was shown to complement the cardiovascular phenotype, demonstrating that indeed haploinsufficiency of one or more genes from this area results in CHD. Other typical *del22q11* features were not found in these animals raising the possibility that additional genes outside the deleted region are involved. Curiously, *Ufd1*, which had previously been implicated in human *del22q11* CHD remained deleted after the complementation experiments without apparent phenotypic consequence (Lindsay et al., 1999; Yamagishi et al., 1999).

Additional haploinsufficient mice generated in parallel with those of Lindsay and co-workers included smaller deletions encompassing several groups of genes from the murine *del22q11* syntenic region. One mouse carried loss of approximately 150 kb from mouse chromosome 16 and included at least seven genes (Kimber et al., 1999). These mice did not show any typical anatomical features of *del22q11* syndrome, nor were there any histological abnormalities. A second group produced mice haploinsufficient for a 550 kb section harboring 16 orthologous genes from the *del22q11* region (Puech et al., 2000). Again, these mice revealed no typical features of *del22q11* syndrome.

Nevertheless, the value of both models became apparent when contrasted with the results of Lindsay and colleagues by narrowing the critical region to several remaining candidate genes.

In follow up studies several groups used separate techniques, including single gene deletion and a series of Cre-*loxP* mediated nested deletions to dissect the remaining critical region. Both groups found that mice lacking a copy of the T-box transcription factor *Tbx1* show typical cardiovascular defects associated with *del22q11* syndrome, as well as other associated phenotypic anomalies (Jerome and Papaioannou, 2001; Lindsay et al., 2001). Furthermore, Lindsay et al. was able to complement *Tbx1* mutations by breeding $Tbx1^{+/-}$ mutants with transgenic mice containing an integrated P1 artificial chromosome known to include a copy of *Tbx1*.

In an impressively exhaustive study, Merscher and co-workers generated mice shown to harbor deletion of a 24-gene region on mouse chromosome 16 (*Lgdel/+* mice) (Merscher et al., 2001). These haploinsufficient mice showed conotruncal anomalies and VSD similar to those often seen in *del22q11* syndrome. Further, these mice showed parathyroid aplasia, an additional feature typically associated with the human condition. In addition to generating the *Lgdel/+* mice, three additional strains of transgenic mice were produced using non-overlapping human BACs from the region for the purpose of complementing the defects seen in hemizygous animals. Surprisingly, mice overexpressing human *Tbx1* were themselves found to contain cardiovascular defects including VSD and OT abnormalities, as well as thymic hypoplasia demonstrating the dosage sensitivity of this transcription factor. When *Lgdel/+* mice were bred with transgenics overexpressing *Tbx1*, partial rescue of the vascular malformation was seen.

This was not found in experiments where *Lgdel/+* mice were bred with transgenics containing BACs from other areas of the *del22q11* region, implicating *Tbx1* in cardiovascular defects. Finally, to definitively show *Tbx1* is the gene primarily responsible for the phenotype, *Tbx1/+* mutants were made and assessed for CHD. As expected, haploinsufficient mice showed heart defects whereas wild-type mice did not, confirming the previous observations (Merscher et al., 2001).

Identification of the principal gene responsible for the *del22q11* cardiovascular phenotype was made possible primarily through the use of targeted gene deletion in mice, once again demonstrating the utility of this approach. However, complete recapitulation of the *del22q11* phenotype was not found in any of the mutant animals discussed above. This highlights the major limitation of applying this technique; that is, there are often fundamental and insurmountable differences in biology between species. Nevertheless, even these limits can yield valuable information when considered in a broader context. For example, in no case did heterozygous *del22q11* completely mimic the human disorder. However, a thorough examination of $Tbx1^{-/-}$ embryos discovered diverse phenotypic consequences, including OT anomalies, thymic aplasia, abnormal facial structures and cleft palate as well as under-formed vertebrae and mandible (Jerome and Papaioannou, 2001). Based on these findings, it was postulated that perhaps humans are more sensitive to dosage effects of this transcription factor and thus suffer a wider array of organ and tissue involvement as a consequence of reduced *TBX1* expression. Further, because it has been shown that non-overlapping deletions in humans (including deletions that exclude *TBX1*) can give rise to a similar phenotype, it has been proposed that regulatory elements or modifiers in the deleted areas (and outside coding regions) are

partly responsible for *del22q11* syndrome (Novelli et al., 1999). Moreover, it has been suggested that other genes from the commonly deleted region exert their own effects when hemizygous only when they occur in a background of *Tbx1* haploinsufficiency (Jerome and Papaioannou, 2001). These suggestions are likely to be tested further in additional models, such as double-heterozygous mice with different combinations of regions or single candidate genes in conjunction with *Tbx1* knocked out.

Parallels exist between *del22q11* syndrome and 3p- syndrome. These primarily include broad phenotypic variability associated with heterogeneous loss of multiple genes. Furthermore, the majority of cases arise *de novo* with haploinsufficiency the likely mechanism. These recurring rearrangements are probably due to regional sequence characteristics that predispose these regions to instability (Stankiewicz and Lupski, 2002). In fact, the presence of deletion 'clusters' within the genome supports this idea (Brewer et al., 1998). However, while *Tbx1* mutations have been implicated in CHD associated with *del22q11*, the gene(s) responsible for AVSD in 3p- syndrome are yet undefined.

The establishment of a *Creld1*$^{+/-}$ colony affords the opportunity to investigate the role of this gene in the etiology of CHD associated with 3p- syndrome. This approach was taken after early mapping studies positioned *CRELD1* within the CHD critical region. Since initiating the task of producing *Creld1* knockout mice, additional mapping studies have been conducted.

In a follow up study of ten 3p- syndrome patients (five of whom were CHD positive) Green and colleagues performed detailed mapping with polymorphic markers and further refined the CHD critical region to markers D3S1263 (proximal) and D3S3594

(distal). Based on these data, *ATP2B2* (OMIM#108733), fibulin-2 (OMIM#135821), *TIMP4* (OMIM#601915) and *SEC13R* (OMIM#600152) were eliminated from consideration as CHD candidates (Green et al., 2000). Interestingly, the narrowed critical region has recently been designated atrioventricular septal defect 2 (AVSD2; OMIM#606217), and according to the latest iteration of the draft sequence of the human genome, *CRELD1* is positioned telomeric to this gene cluster. If these results are correct, *CRELD1* should also be excluded from consideration as a CHD candidate gene. However, this conclusion is predicated on the notions that the phenotype is fully penetrant and the mapping data is complete.

Though lesions resulting in haploinsufficiency for both *NKX2-5* and *JAG1* have been shown to cause both syndromic and non-syndromic CHD (Li et al., 1997a; Schott et al., 1998), mutations in these genes are not completely penetrant (Eldadah et al., 2001; Goldmuntz et al., 2001). This raises the possibility that haploinsufficiency of one or more genes involved in heart development and deleted in 3p- syndrome might yield an incompletely penetrant phenotype as well. If this were true, it is possible that while the most proximal break point firmly establishes a centromeric boundary for the CHD critical region, delineation of the distal boundary would be more difficult by mapping data alone.

In addition to issues of incomplete penetrance, close examination of regional microsatellite marker placement available through public databases demonstrates uncertainty, and our most recent search (7/10/02) for one of the two published CHD anchor loci (D3S1263) showed the marker was positioned in two separate locations on distal chromosome 3p25. (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/map_search). The interval between the alternative map positions for D3S1263 varies over several Mb

(compare Whitehead Institute YAC and RH maps with Genethon, Marshfield and STS maps). *CRELD1* is localized to this region of ambiguity. While these findings must be treated with care, they do not exclude the possibility that refinement in mapping data and sequence annotation may adjust the critical region boundaries.

Recent reports stress the intractable characteristics of certain chromosomal regions that make cloning, sequence production and assembly difficult. Further, though sequencing of several chromosomes (20, 21, 22, and Y) is reported to be virtually complete, annotated sequence of chromosome 3 is only approximately 50% finished (Pennisi, 2002). One of the factors exacerbating the automated assembly of any contiguous sequence is the presence of repetitive elements, encountered frequently near the telomeres. During the cloning of mouse *Creld1* multiple repetitive sequences were identified (SINE and LINE class), demonstrating the presence of these low complexity repeats in mice. Similarly, analysis of *CRELD1* genomic sequence showed these repeats also exist in human. Presently, it is unclear if this is true for large regions of distal chromosome 3, though this is a possible explanation for the uncertainty in the genetic map across this region.

The most recent paper brought to bear on the genetics of 3p- syndrome describes a subtle interstitial deletion of chromosome 3p25.3-26.2 (Cargile et al., 2002). FISH analysis using STS-linked BAC clones placed the deleted region between D3S3630 and D3S1304. This region spans approximately 4.5 Mb and is the smallest 3p- syndrome deletion reported to date. This study is interesting in that the patient presents with typical 3p- syndrome characteristics, though lacks CHD. According to the currently available data, *CRELD1* lies centromeric of this deletion towards the CHD critical region. A

particularly intriguing finding is that the relative marker placement across the CHD critical region originally presented by Green et al. was recently reported to be different by Cargile and co-workers, demonstrating the changing nature of the mapping data in this area (Green et al., 2000; Cargile et al., 2002). Bearing these factors in mind, *CRELD1* cannot yet be conclusively excluded from consideration in the etiology of congenital heart defects in 3p- syndrome.

Thus, our immediate goals will focus on phenotypic assessment of the newly generated *Creld1*$^{+/-}$ mice. Thus far, sufficient offspring have been generated to both propagate the colony as well as conduct gross anatomical studies and *Creld1* heterozygous mutants appear normal. However, in the course of colony expansion, a deviation from Mendelian inheritance was ascertained. Timed pregnancy, embryo recovery and genotyping have demonstrated *Creld1* homozygous-null mice die *in utero*. Embryos at different gestational timepoints have been obtained (e9.5-13.5). It appears that death is occurring early in development, since no *Creld1*$^{-/-}$ embryos can be identified beyond e12.5. Given that the heart is the first functional organ to develop (~e9.0), it is possible that the primary cause of embryonic lethality in these mice is an underlying cardiac defect. More data needs to be collected to prove (or disprove) this hypothesis.

Meanwhile, the anatomy of the *Creld1*$^{+/-}$ mouse cardiovascular system will be examined in detail. Gross appearance of excised hearts demonstrating cardiac malformation has been previously observed in haploinsufficient mouse models. A particularly elegant assessment of heart morphology was presented for the *Tbx5*$^{+/-}$ mouse (Bruneau et al., 2001). Whole hearts were subject to MRI analysis to visualize ASD and dilation of the right atrium and ventricle in the intact organ. This study also relied on

standard methodology, such as the histological examination of serial sections from mutant and wild-type littermates. Additionally, ECG analysis was performed on these mice to detect the physiological consequences of the introduced mutation. ECG technology for mouse examination is currently available on the OHSU campus, and it will be important to bear in mind that while *Creld1*$^{+/-}$ mice thus far appear normal, a more complete assessment may show subtle changes in heart physiology as well as anatomy.

Other methods for examination of heart morphology have been reported, though they appear more challenging and likely require substantial expertise. For example, Lindsay and colleagues illustrate changes in the pharyngeal arch arteries in heterozygous mutant *Tbx1* animals by intracardiac India-ink injection. Their technique relied on collecting embryos followed by injection and rapid fixing in a solution of PBS-buffered formaldehyde (Lindsay et al., 2001). A second means by which heart anatomy can be detailed is described by Merscher and co-workers (Merscher et al., 2001). This group was also studying *Tbx1* haploinsufficiency and were able to detail defective conotruncal development by making corrosion casts via embryonic methylmethacrilate injection followed by dissolution of soft tissue with commercially available maceration solution. This method was also useful in detecting a VSD as the resin was observed to quickly and easily cross the ventricular septum. Again, these methods are technically sophisticated and would probably require outside assistance.

As mentioned previously, *CRELD1* is expressed not only in heart, but also in other tissues during development such as brain, liver, kidney, lung and limb buds (Rupp, Fouad et al., 2002). This raised the possibility that non-cardiac tissues may be affected

by our targeted mutation. Again, while data collection is ongoing, qualitatively the *Creld1*-null embryos show widespread tissue underdevelopment. Whether or not this results from insufficient cardiac output (necessary for rapidly growing tissues) or from a primary defect in these tissues themselves remains to be determined.

To date, no significant malformations in *Creld1*$^{+/-}$ mice have been identified. However, there remains the possibility that slight or inconspicuous defects may be present. Furthermore, it is possible that an underlying physiological defect may emerge with increasing age. Therefore, in order to reduce the chances of overlooking subtle developmental changes in *Creld1*$^{+/-}$ mice, we will also pay close attention to the anatomy of other areas, specifically limb as well as skeletal and nervous systems. Alcian blue and Alizarin red to assess changes in bone and cartilage have been described and could prove useful in assessment of *Creld1*$^{+/-}$ mice (Jerome and Papaioannou, 2001).

Finally, because mouse models have been produced for both SPARC and decorin, it is conceivable that double mutants could be generated. Both the SPARC and decorin mouse phenotypes were reported to be relatively mild, given the apparent widespread involvement of these molecules during development (Danielson et al., 1997; Gilmour et al., 1998; Norose et al., 1998). A double mutant demonstrating a more dramatic phenotype would likely add insight not only regarding the *in vivo* biological role of CRELD1, but those of SPARC and decorin as well.

Conclusion

This work describes the production of a *Creld1* knockout mouse. While *Creld1* haploinsufficient mice appear normal, it is now clear that completely abolishing *Creld1* expression is not compatible with survival. As mentioned, colony expansion is currently underway and assessment of heart morphology is our highest priority in both heterozygous and (embryonic) homozygous mutants. We previously hypothesized that *CRELD1* haploinsufficiency may lead to AVSD based on early studies that defined a region at chromosome 3p25-pter essential for normal cardiac development. Deletion of this region has been shown to cause 3p- syndrome in humans. More proximal deletions associated with 3p- syndrome have been correlated with congenital heart defects, though the specific gene or genes responsible for these observations are not yet known.

CRELD1 was initially localized to the CHD critical region associated with 3p- syndrome, and while recent work demonstrates some confusion with regard to the proper placement of *CRELD1* relative to markers that define the critical region, we believe *CRELD1* makes a compelling candidate gene for involvement in heart development based on its spatial and temporal patterns of expression. Strong *CRELD1* expression in heart has been demonstrated by whole mount *in situ* hybridization, on fetal and adult multiple tissue Northern blots as well as in a multiple tissue heart-specific Northern blot. Furthermore, a related project in our laboratory has identified several putative mutations in ethnically diverse individuals with cardiac malformation. This is perhaps the strongest evidence yet that *CRELD1* is essential to proper cardiogenesis.

While the primary goal of generating *Creld1* knockout mice was to test the hypothesis that haploinsufficiency of *Creld1* causes CHD, simple recapitulation of AVSD typically associated with more severely affected 3p- syndrome patients, though informative, does not exhaust the value of having such an animal model. It is our hope that having a murine model defective in atrioventricular septation, in combination with other materials and techniques, will be a useful tool in further elucidating the series of events leading to appropriate chamber formation during the process of heart development.

Because of the unique sequence segments found in *CRELD1*, prediction of function based on similarity with other proteins is difficult. However, the presence of modules typically associated with extracellular matrix molecules and a predicted transmembrane spanning domain in the amino acid sequence has lead us to consider the possibility that CRELD1 is involved in the process of cell adhesion. We have conducted a yeast two-hybrid screen using the highly conserved CRELD1 WE domain. By this method we identified two CRELD1 binding partners: SPARC and decorin. These two proteins are small multifunctional extracellular matrix molecules that have been implicated in processes that include regulating cell proliferation, cellular de-adhesion and migration. In the future, it will be important to further characterize these interactions with the long term aim of identifying functional pathways regulated by these proteins. It is hoped that this work will provide the foundation upon which these goals are attained.

References

Alagille, D., Odievre, M., Gautier, M. and Dommergues, J.P.: Hepatic ductular hypoplasia associated with characteristic facies, vertebral malformations, retarded physical, mental, and sexual development, and cardiac murmur. J Pediatr 86 (1975) 63-71.

Artavanis-Tsakonas, S., Matsuno, K. and Fortini, M.E.: Notch signaling. Science 268 (1995) 225-32.

Asano, K., Vornlocher, H.P., Richter-Cook, N.J., Merrick, W.C., Hinnebusch, A.G. and Hershey, J.W.: Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits. Possible roles in RNA binding and macromolecular assembly. J Biol Chem 272 (1997) 27042-52.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K.: Current Protocols In Molecular Biology. In: Chanda, V.B. (Ed.). John Wiley & Sons, Inc., 2001.

Baldini, A.: DiGeorge syndrome: the use of model organisms to dissect complex genetics. Hum Mol Genet 11(2002) 2363-2369.

Barlow, G.M., Chen, X.N., Shi, Z.Y., Lyons, G.E., Kurnit, D.M., Celle, L., Spinner, N.B., Zackai, E., Pettenati, M.J., Van Riper, A.J., Vekemans, M.J., Mjaatvedt, C.H., and Korenberg, J.R.: Down syndrome congenital heart disease: A narrowed region and a candidate gene. Genet Med 3 (2001) 91-101.

Basson, C.T., Bachinsky, D.R., Lin, R.C., Levi, T., Elkins, J.A., Soults, J., Grayzel, D., Kroumpouzou, E., Traill, T.A., Leblanc-Straceski, J., Renault, B., Kucherlapati, R., Seidman, J.G. and Seidman, C.E.: Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. Nat Genet 15 (1997) 30-5.

Bhavsar, P.K., Brand, N.J., Yacoub, M.H. and Barton, P.J.: Isolation and characterization of the human cardiac troponin I gene (TNNI3). Genomics 35 (1996) 11-23.

Bisaha, J.G. and Bader, D.: Identification and characterization of a ventricular-specific avian myosin heavy chain, VMHC1: expression in differentiating cardiac and skeletal muscle. Dev Biol 148 (1991) 355-64.

Blake, J.A., Eppig, J.T., Richardson, J.E., Bult, C.J. and Kadin, J.A.: The Mouse Genome Database (MGD): integration nexus for the laboratory mouse. Nucleic Acids Res 29 (2001) 91-4.

Bodmer, R.: The gene tinman is required for specification of the heart and visceral muscles in Drosophila. Development 118 (1993) 719-29.

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

Bouchey, D., Argraves, W.S. and Little, C.D.: Fibulin-1, vitronectin, and fibronectin expression during avian cardiac valve and septa development. Anat Rec 244 (1996) 540-51.

Boyer, B., Valles, A.M. and Edme, N.: Induction and regulation of epithelial-mesenchymal transitions. Biochem Pharmacol 60 (2000) 1091-9.

Bradshaw, A.D., Francki, A., Motamed, K., Howe, C. and Sage, E.H.: Primary mesenchymal cells isolated from SPARC-null mice exhibit altered morphology and rates of proliferation. Mol Biol Cell 10 (1999) 1569-79.

Bradshaw, A.D. and Sage, E.H.: SPARC, a matricellular protein that functions in cellular differentiation and tissue response to injury. J Clin Invest 107 (2001) 1049-54.

Brekken, R.A. and Sage, E.H.: SPARC, a matricellular protein: at the crossroads of cell-matrix communication. Matrix Biol 19 (2001) 816-27.

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

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

Bunting, M., Bernstein, K.E., Greer, J.M., Capecchi, M.R. and Thomas, K.R.: Targeting genes for self-excision in the germ line. Genes Dev 13 (1999) 1524-8.

Cargile, C.B., Goh, D.L., Goodman, B.K., Chen, X.N., Korenberg, J.R., Semenza, G.L. and Thomas, G.H.: Molecular cytogenetic characterization of a subtle interstitial del(3)(p25.3p26.2) in a patient with deletion 3p syndrome. Am J Med Genet 109 (2002) 133-8.

Coffin, J.D. and Poole, T.J.: Embryonic vascular development: immunohistochemical identification of the origin and subsequent morphogenesis of the major vessel primordia in quail embryos. Development 102 (1988) 735-48.

Cohen-Gould, L. and Mikawa, T.: The fate diversity of mesodermal cells within the heart field during chicken early embryogenesis. Dev Biol 177 (1996) 265-73.

Colomer, V., Engelender, S., Sharp, A.H., Duan, K., Cooper, J.K., Lanahan, A., Lyford, G., Worley, P. and Ross, C.A.: Huntingtin-associated protein 1 (HAP1) binds to a Trio-like polypeptide, with a rac1 guanine nucleotide exchange factor domain. Hum Mol Genet 6 (1997) 1519-25.

Corbin, V., Michelson, A.M., Abmayr, S.M., Neel, V., Alcamo, E., Maniatis, T. and Young, M.W.: A role for the Drosophila neurogenic genes in mesoderm differentiation. Cell 67 (1991) 311-23.

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

Danielson, K.G., Baribault, H., Holmes, D.F., Graham, H., Kadler, K.E. and Iozzo, R.V.: Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. J Cell Biol 136 (1997) 729-43.

Danielson, K.G., Fazzio, A., Cohen, I., Cannizzaro, L.A., Eichstetter, I. and Iozzo, R.V.: The human decorin gene: intron-exon organization, discovery of two alternatively spliced exons in the 5' untranslated region, and mapping of the gene to chromosome 12q23. Genomics 15 (1993) 146-60.

Davies C.D.L., Melder, R.J., Munn, L.L., Mouta-Carreira, C., Jain, R.K. and Boucher, Y.: Decorin inhibits endothelial migration and tube-like structure formation: role of thrombospondin-1. Microvasc Res 62 (2001) 26-42.

Davis, C.L.: The cardiac jelly of the chick embryo. Anat Rec 27 (1924) 201-202.

Davis, C.L.: Development of the human heart from its first appearance to the stage found in embryos of twenty paired somites. Contributions to Embryology 19 (1927) 245-284.

Davis, G.E. and Camarillo, C.W.: An alpha 2 beta 1 integrin-dependent pinocytic mechanism involving intracellular vacuole formation and coalescence regulates capillary lumen and tube formation in three-dimensional collagen matrix. Exp Cell Res 224 (1996) 39-51.

de la Cruz, M., Sanchez-Gomez, C., and Palomino, M.: The primitive cardiac regions in the straight tube heart (Stage9-) and their anatomical expression in the mature heart: an experimental study in chick embryo. Journal of Anatomy 165 (1989) 121-131.

Delany, A.M., Amling, M., Priemel, M., Howe, C., Baron, R. and Canalis, E.: Osteopenia and decreased bone formation in osteonectin-deficient mice. J Clin Invest 105 (2000) 1325.

Dietz, H.C., Cutting, G.R., Pyeritz, R.E., Maslen, C.L., Sakai, L.Y., Corson, G.M., Puffenberger, E.G., Hamosh, A., Nanthakumar, E.J., Curristin, S.M. and et al.: Marfan

syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature 352 (1991) 337-9.

Drumheller, T., McGillivray, B.C., Behrner, D., MacLeod, P., McFadden, D.E., Roberson, J., Venditti, C., Chorney, K., Chorney, M. and Smith, D.I.: Precise localisation of 3p25 breakpoints in four patients with the 3p-syndrome. J Med Genet 33 (1996) 842-7.

Dunlevy, J.R. and Couchman, J.R.: Controlled induction of focal adhesion disassembly and migration in primary fibroblasts. J Cell Sci 105 (1993) 489-500.

Dunlevy, J.R. and Couchman, J.R.: Interleukin-8 induces motile behavior and loss of focal adhesions in primary fibroblasts. J Cell Sci 108 (1995) 311-21.

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

Eldadah, Z.A., Hamosh, A., Biery, N.J., Montgomery, R.A., Duke, M., Elkins, R. and Dietz, H.C.: Familial Tetralogy of Fallot caused by mutation in the jagged1 gene. Hum Mol Genet 10 (2001) 163-9.

Epstein, J.A.: Developing models of DiGeorge syndrome. Trends Genet 17 (2001) S13-7.

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

Farrell, M., Waldo, K., Li, Y.X. and Kirby, M.L.: A novel role for cardiac neural crest in heart development. Trends Cardiovasc Med 9 (1999) 214-20.

Firulli, A.B., McFadden, D.G., Lin, Q., Srivastava, D. and Olson, E.N.: Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. Nat Genet 18 (1998) 266-70.

Fishman, M.C. and Chien, K.R.: Fashioning the vertebrate heart: earliest embryonic decisions. Development 124 (1997) 2099-117.

Fishman, M.C. and Olson, E.N.: Parsing the heart: genetic modules for organ assembly. Cell 91 (1997) 153-6.

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

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

Gassmann, M., Casagranda, F., Orioli, D., Simon, H., Lai, C., Klein, R. and Lemke, G.: Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. Nature 378 (1995) 390-4.

Giancotti, F.G. and Ruoslahti, E.: Integrin signaling. Science 285 (1999) 1028-32.
Gilbert, S.F.: Developmental Biology, 6th ed. Sinauer Associates, Sunderland, MA, 2000.

Gilmour, D.T., Lyon, G.J., Carlton, M.B., Sanes, J.R., Cunningham, J.M., Anderson, J.R., Hogan, B.L., Evans, M.J. and Colledge, W.H.: Mice deficient for the secreted glycoprotein SPARC/osteonectin/BM40 develop normally but show severe age-onset cataract formation and disruption of the lens. Embo J 17 (1998) 1860-70.

Goldblum, S.E., Ding, X., Funk, S.E. and Sage, E.H.: SPARC (secreted protein acidic and rich in cysteine) regulates endothelial cell shape and barrier function. Proc Natl Acad Sci U S A 91 (1994) 3448-52.

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

Gottlieb, P.D., Pierce, S.A., Sims, R.J., Yamagishi, H., Weihe, E.K., Harriss, J.V., Maika, S.D., Kuziel, W.A., King, H.L., Olson, E.N., Nakagawa, O. and Srivastava, D.: Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. Nat Genet 31 (2002) 25-32.

Grech, V. and Gatt, M.: Syndromes and malformations associated with congenital heart disease in a population-based study. Int J Cardiol 68 (1999) 151-6.

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

Harrison, P.M. and Arosio, P.: The ferritins: molecular properties, iron storage function and cellular regulation. Biochim Biophys Acta 1275 (1996) 161-203.

Harvey, R.P.: NK-2 homeobox genes and heart development. Dev Biol 178 (1996) 203-16.

Hausser, H., Schonherr, E., Muller, M., Liszio, C., Bin, Z., Fisher, L.W. and Kresse, H.: Receptor-mediated endocytosis of decorin: involvement of leucine-rich repeat structures. Arch Biochem Biophys 349 (1998) 363-70.

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

Hoch, M., Broadie, K., Jackle, H. and Skaer, H.: Sequential fates in a single cell are established by the neurogenic cascade in the Malpighian tubules of Drosophila. Development 120 (1994) 3439-50.

Hoffman, J.I.: Incidence of congenital heart disease: I. Postnatal incidence. Pediatr Cardiol 16 (1995a) 103-13.

Hoffman, J.I.: Incidence of congenital heart disease: II. Prenatal incidence. Pediatr Cardiol 16 (1995b) 155-65.

Hofmann, K. and Bucher, P.: The PCI domain: a common theme in three multiprotein complexes. Trends Biochem Sci 23 (1998) 204-5.

Hohenester, E., Maurer, P., Hohenadl, C., Timpl, R., Jansonius, J.N. and Engel, J.: Structure of a novel extracellular Ca(2+)-binding module in BM-40. Nat Struct Biol 3 (1996) 67-73.

Hohenester, E., Maurer, P. and Timpl, R.: Crystal structure of a pair of follistatin-like and EF-hand calcium-binding domains in BM-40. Embo J 16 (1997) 3778-86.

Hynes, R.O. and Zhao, Q.: The evolution of cell adhesion. J Cell Biol 150 (2000) F89-96.

Iozzo, R.V.: The biology of the small leucine-rich proteoglycans. Functional network of interactive proteins. J Biol Chem 274 (1999) 18843-6.

Iozzo, R.V., Chakrani, F., Perrotti, D., McQuillan, D.J., Skorski, T., Calabretta, B. and Eichstetter, I.: Cooperative action of germ-line mutations in decorin and p53 accelerates lymphoma tumorigenesis. Proc Natl Acad Sci U S A 96 (1999a) 3092-7.

Iozzo, R.V., Moscatello, D.K., McQuillan, D.J. and Eichstetter, I.: Decorin is a biological ligand for the epidermal growth factor receptor. J Biol Chem 274 (1999b) 4489-92.

Jerome, L.A. and Papaioannou, V.E.: DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. Nat Genet 27 (2001) 286-91.

Johnson, J.S., Silberbach, G.M., Brown, M.G., Lawce, H., Olson, S.B., and Pillers, D.M. Screening neonates with congenital heart disease for 22q11.2 deletion. J Invest Med (2002) 50(1):19A.

Kiemer, A.K., Takeuchi, K. and Quinlan, M.P.: Identification of genes involved in epithelial-mesenchymal transition and tumor progression. Oncogene 20 (2001) 6679-88.

Kimber, W.L., Hsieh, P., Hirotsune, S., Yuva-Paylor, L., Sutherland, H.F., Chen, A., Ruiz-Lozano, P., Hoogstraten-Miller, S.L., Chien, K.R., Paylor, R., Scambler, P.J. and

Wynshaw-Boris, A.: Deletion of 150 kb in the minimal DiGeorge/velocardiofacial syndrome critical region in mouse. Hum Mol Genet 8 (1999) 2229-37.

Kinsella, M.G., Fischer, J.W., Mason, D.P. and Wight, T.N.: Retrovirally mediated expression of decorin by macrovascular endothelial cells. Effects on cellular migration and fibronectin fibrillogenesis in vitro. J Biol Chem 275 (2000) 13924-32.
Kirby, M.L. and Waldo, K.L.: Neural crest and cardiovascular patterning. Circ Res 77 (1995) 211-5.

Korenberg, J.R., Chen, X.N., Schipper, R., Sun, Z., Gonsky, R., Gerwehr, S., Carpenter, N., Daumer, C., Dignan, P., Disteche, C., Graham, J.M. Jr., Hugdins, L., McGillivray, B., Miyazaki, K., Ogasawara, N., Park, J.P., Pagon, R., Pueschel, S., Sack, G., Say, B., Schuffenhauer, S., Soukup, S., and Yamanaka, T.: Down syndrome phenotypes: The consequences of chromosomal imbalance. Proc Natl Acad Sci U S A 91 (1994) 4997-5001.

Korenberg, J.R., Kawashima, H., Pulst, S.M., Ikeuchi, T., Ogasawara, N., Yamamoto, K., Schonberg, S.A., West, R., Allen, L., Magenis, E., Ikawa, K., Taniguchi, N., and Epstein, C.J.: Molecular definition of a region of chromosome 21 that causes features of the Down syndrome phenotype. Am J Hum Genet 47 (1990) 236-46.

Krug, E.L., Mjaatvedt, C.H. and Markwald, R.R.: Extracellular matrix from embryonic myocardium elicits an early morphogenetic event in cardiac endothelial differentiation. Dev Biol 120 (1987) 348-55.

Kubalak, S.W., Miller-Hance, W.C., O'Brien, T.X., Dyson, E. and Chien, K.R.: Chamber specification of atrial myosin light chain-2 expression precedes septation during murine cardiogenesis. J Biol Chem 269 (1994) 16961-70.

Kuo, C.T., Morrisey, E.E., Anandappa, R., Sigrist, K., Lu, M.M., Parmacek, M.S., Soudais, C. and Leiden, J.M.: GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Genes Dev 11 (1997) 1048-60.

Lamers, W.H., Geerts, W.J. and Moorman, A.F.: Distribution pattern of acetylcholinesterase in early embryonic chicken hearts. Anat Rec 228 (1990) 297-305.

Lamers, W.H., te Kortschot, A., Los, J.A. and Moorman, A.F.: Acetylcholinesterase in prenatal rat heart: a marker for the early development of the cardiac conductive tissue? Anat Rec 217 (1987) 361-70.

Lane, T.F., Iruela-Arispe, M.L. and Sage, E.H.: Regulation of gene expression by SPARC during angiogenesis in vitro. Changes in fibronectin, thrombospondin-1, and plasminogen activator inhibitor-1. J Biol Chem 267 (1992) 16736-45.

Ledda, M.F., Adris, S., Bravo, A.I., Kairiyama, C., Bover, L., Chernajovsky, Y., Mordoh, J. and Podhajcer, O.L.: Suppression of SPARC expression by antisense RNA abrogates the tumorigenicity of human melanoma cells. Nat Med 3 (1997) 171-6.

Li, B. and Fields, S.: Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. Faseb J 7 (1993) 957-63.

Li, L., Krantz, I.D., Deng, Y., Genin, A., Banta, A.B., Collins, C.C., Qi, M., Trask, B.J., Kuo, W.L., Cochran, J., Costa, T., Pierpont, M.E., Rand, E.B., Piccoli, D.A., Hood, L. and Spinner, N.B.: Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. Nat Genet 16 (1997a) 243-51.

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

Lin, Q., Schwarz, J., Bucana, C. and Olson, E.N.: Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. Science 276 (1997) 1404-7.

Linask, K.K., Knudsen, K.A. and Gui, Y.H.: N-cadherin-catenin interaction: necessary component of cardiac cell compartmentalization during early vertebrate heart development. Dev Biol 185 (1997) 148-64.

Linask, K.K. and Lash, J.W.: A role for fibronectin in the migration of avian precardiac cells. I. Dose-dependent effects of fibronectin antibody. Dev Biol 129 (1988a) 315-23.

Linask, K.K. and Lash, J.W.: A role for fibronectin in the migration of avian precardiac cells. II. Rotation of the heart-forming region during different stages and its effects. Dev Biol 129 (1988b) 324-9.

Lindsay, E.A.: Chromosomal microdeletions: dissecting del22q11 syndrome. Nat Rev Genet 2 (2001) 858-68.

Lindsay, E.A., Botta, A., Jurecic, V., Carattini-Rivera, S., Cheah, Y.C., Rosenblatt, H.M., Bradley, A. and Baldini, A.: Congenital heart disease in mice deficient for the DiGeorge syndrome region. Nature 401 (1999) 379-83.

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

Lindsell, C.E., Shawber, C.J., Boulter, J. and Weinmaster, G.: Jagged: a mammalian ligand that activates Notch1. Cell 80 (1995) 909-17.

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

Manner, J.: Cardiac looping in the chick embryo: a morphological review with special reference to terminological and biomechanical aspects of the looping process. Anat Rec 259 (2000) 248-62.

Maslen, C.L. and Glanville, R.W..: The Marfan Syndrome. The Endocrinologist 3 (1993) 279-287.

Maslen, C.L.: Fundamental advances in molecular genetics: new insights and detection strategies in congenital heart disease, Congenital Heart Disease in the Adult. American College of Cardiology Educational Highlights Vol 12 (1997) 11-14 .

Maslen, C.L., Corson, G.M., Maddox, B.K., Glanville, R.W. and Sakai, L.Y.: Partial sequence of a candidate gene for the Marfan syndrome. Nature 352 (1991) 334-7.

Maslen, C.L., Rupp, P.A., Olson, S.B., Reifsteck, C.A., Thornburg, K.L., and Glanville, R.W.: Characterization of cirrin, a new extracellular protein gene that is associated with endocardial cushion defects in 3p- syndrome. Am J Hum Genet 65 (1999) A43.

Masson, S., Arosio, B., Luvara, G., Gagliano, N., Fiordaliso, F., Santambrogio, D., Vergani, C., Latini, R. and Annoni, G.: Remodelling of cardiac extracellular matrix during beta-adrenergic stimulation: upregulation of SPARC in the myocardium of adult rats. J Mol Cell Cardiol 30 (1998) 1505-14.

McCarthy, L.C., Terrett, J., Davis, M.E., Knights, C.J., Smith, A.L., Critcher, R., Schmitt, K., Hudson, J., Spurr, N.K. and Goodfellow, P.N.: A first-generation whole genome-radiation hybrid map spanning the mouse genome. Genome Res 7 (1997) 1153-61.

Merscher, S., Funke, B., Epstein, J.A., Heyer, J., Puech, A., Lu, M.M., Xavier, R.J., Demay, M.B., Russell, R.G., Factor, S., Tokooya, K., Jore, B.S., Lopez, M., Pandita, R.K., Lia, M., Carrion, D., Xu, H., Schorle, H., Kobler, J.B., Scambler, P., Wynshaw-Boris, A., Skoultchi, A.I., Morrow, B.E. and Kucherlapati, R.: TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. Cell 104 (2001) 619-29.

Meyer, D. and Birchmeier, C.: Multiple essential functions of neuregulin in development. Nature 378 (1995) 386-90.

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

Mjaatvedt, C.H., Lepera, R.C. and Markwald, R.R.: Myocardial specificity for initiating endothelial-mesenchymal cell transition in embryonic chick heart correlates with a particulate distribution of fibronectin. Dev Biol 119 (1987) 59-67.

Mjaatvedt, C.H. and Markwald, R.R.: Induction of an epithelial-mesenchymal transition by an in vivo adheron-like complex. Dev Biol 136 (1989) 118-28.

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

Mjaatvedt, C.H., Yamamura, H., Wessels, A., Ramsdell, A., Turner, D and Markwald, R.R.: Mechanisms of Segmentation, Septation, and Remodeling of the Tubular Heart: Endocardial Cushion fate and Cardiac Looping. In: Harvey, R.P., and Rosenthal, N (Ed.), Heart Development. Academic Press, San Diego, 1999, pp. 159-177.

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

Morimura, N., Tezuka, Y., Watanabe, N., Yasuda, M., Miyatani, S., Hozumi, N. and Tezuka Ki, K.: Molecular cloning of POEM: a novel adhesion molecule that interacts with alpha8beta1 integrin. J Biol Chem 276 (2001) 42172-81.

Mowery, P.N., Chorney, M.J., Venditti, C.P., Latif, F., Modi, W.S., Lerman, M.I., Zbar, B., Robins, D.B., Rogan, P.K., and Ladda, R.L.: Clinical and molecular analyses of deletion 3p25-pter syndrome. Am J Med Genet (1993).

Murphy-Ullrich, J.E.: The de-adhesive activity of matricellular proteins: is intermediate cell adhesion an adaptive state? J Clin Invest 107 (2001) 785-90.

Murphy-Ullrich, J.E., Lane, T.F., Pallero, M.A. and Sage, E.H.: SPARC mediates focal adhesion disassembly in endothelial cells through a follistatin-like region and the Ca(2+)-binding EF-hand. J Cell Biochem 57 (1995) 341-50.

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

Nakajima, Y., Yamagishi, T., Nakamura, H., Markwald, R.R. and Krug, E.L.: An autocrine function for transforming growth factor (TGF)-beta3 in the transformation of

atrioventricular canal endocardium into mesenchyme during chick heart development. Dev Biol 194 (1998) 99-113.

Nickerson, E., Greenberg, F., Keating, M.T., McCaskill, C. and Shaffer, L.G.: Deletions of the elastin gene at 7q11.23 occur in approximately 90% of patients with Williams syndrome. Am J Hum Genet 56 (1995) 1156-61.

Norose, K., Clark, J.I., Syed, N.A., Basu, A., Heber-Katz, E., Sage, E.H. and Howe, C.C.: SPARC deficiency leads to early-onset cataractogenesis. Invest Ophthalmol Vis Sci 39 (1998) 2674-80.

Novelli, G., Amati, F. and Dallapiccola, B.: UFD1L and CDC45L: a role in DiGeorge syndrome and related phenotypes? Trends Genet 15 (1999) 251-4.

O'Brien, T.X., Lee, K.J. and Chien, K.R.: Positional specification of ventricular myosin light chain 2 expression in the primitive murine heart tube. Proc Natl Acad Sci U S A 90 (1993) 5157-61.

Oda, T., Elkahloun, A.G., Meltzer, P.S. and Chandrasekharappa, S.C.: Identification and cloning of the human homolog (JAG1) of the rat Jagged1 gene from the Alagille syndrome critical region at 20p12. Genomics 43 (1997a) 376-9.

Oda, T., Elkahloun, A.G., Pike, B.L., Okajima, K., Krantz, I.D., Genin, A., Piccoli, D.A., Meltzer, P.S., Spinner, N.B., Collins, F.S. and Chandrasekharappa, S.C.: Mutations in the human Jagged1 gene are responsible for Alagille syndrome. Nat Genet 16 (1997b) 235-42.

Olson, E.N., Arnold, H.H., Rigby, P.W. and Wold, B.J.: Know your neighbors: three phenotypes in null mutants of the myogenic bHLH gene MRF4. Cell 85 (1996) 1-4.
Olson, E.N. and Srivastava, D.: Molecular pathways controlling heart development. Science 272 (1996) 671-6.

Pennisi, E.: Genomics. Genome centers push for polished draft. Science 296 (2002) 1600-1.

Pereira, L., Andrikopoulos, K., Tian, J., Lee, S.Y., Keene, D.R., Ono, R., Reinhardt, D.P., Sakai, L.Y., Biery, N.J., Bunton, T., Dietz, H.C. and Ramirez, F.: Targeting of the gene encoding fibrillin-1 recapitulates the vascular aspect of Marfan syndrome. Nat Genet 17 (1997) 218-22.

Pereira, L., Lee, S.Y., Gayraud, B., Andrikopoulos, K., Shapiro, S.D., Bunton, T., Biery, N.J., Dietz, H.C., Sakai, L.Y., and Ramirez, F.: Pathogenetic sequence for aneurysm revealed in mice underexpressing fibrillin-1. Proc Natl Acad Sci USA 96 (1999) 3819-23.

Phipps, M.E., Latif, F., Prowse, A., Payne, S.J., Dietz-Band, J., Leversha, M., Affara, N.A., Moore, A.T., Tolmie, J., Schinzel, A. and et al.: Molecular genetic analysis of the 3p- syndrome. Hum Mol Genet 3 (1994) 903-8.

Potts, J.D., Dagle, J.M., Walder, J.A., Weeks, D.L. and Runyan, R.B.: Epithelial-mesenchymal transformation of embryonic cardiac endothelial cells is inhibited by a modified antisense oligodeoxynucleotide to transforming growth factor beta 3. Proc Natl Acad Sci U S A 88 (1991) 1516-20.

Puech, A., Saint-Jore, B., Merscher, S., Russell, R.G., Cherif, D., Sirotkin, H., Xu, H., Factor, S., Kucherlapati, R. and Skoultchi, A.I.: Normal cardiovascular development in mice deficient for 16 genes in 550 kb of the velocardiofacial/DiGeorge syndrome region. Proc Natl Acad Sci U S A 97 (2000) 10090-5.

Radice, G.L., Rayburn, H., Matsunami, H., Knudsen, K.A., Takeichi, M. and Hynes, R.O.: Developmental defects in mouse embryos lacking N-cadherin. Dev Biol 181 (1997) 64-78.

Ramirez-Solis, R., Rivera-Perez, J., Wallace, J.D., Wims, M., Zheng, H. and Bradley, A.: Genomic DNA microextraction: a method to screen numerous samples. Anal Biochem 201 (1992) 331-5.

Reaume, A.G., de Sousa, P.A., Kulkarni, S., Langille, B.L., Zhu, D., Davies, T.C., Juneja, S.C., Kidder, G.M. and Rossant, J.: Cardiac malformation in neonatal mice lacking connexin43. Science 267 (1995) 1831-4.

Reijo, R., Lee, T.Y., Salo, P., Alagappan, R., Brown, L.G., Rosenberg, M., Rozen, S., Jaffe, T., Straus, D., Hovatta, O. and et al.: Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. Nat Genet 10 (1995) 383-93.

Reinboth, B., Hanssen, E., Cleary, E.G. and Gibson, M.A.: Molecular interactions of biglycan and decorin with elastic fiber components: biglycan forms a ternary complex with tropoelastin and microfibril-associated glycoprotein 1. J Biol Chem 277 (2002) 3950-7.

Rempel, S.A., Ge, S. and Gutierrez, J.A.: SPARC: a potential diagnostic marker of invasive meningiomas. Clin Cancer Res 5 (1999) 237-41.

Rezaee, M., Isokawa, K., Halligan, N., Markwald, R.R. and Krug, E.L.: Identification of an extracellular 130-kDa protein involved in early cardiac morphogenesis. J Biol Chem 268 (1993) 14404-11.

Riley, P., Anson-Cartwright, L. and Cross, J.C.: The Hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. Nat Genet 18 (1998) 271-5.

Roberds, S.L., Anderson, R.D., Ibraghimov-Beskrovnaya, O. and Campbell, K.P.: Primary structure and muscle-specific expression of the 50-kDa dystrophin-associated glycoprotein (adhalin). J Biol Chem 268 (1993) 23739-42.

Rosenberg, L.C., Choi, H.U., Tang, L.H., Johnson, T.L., Pal, S., Webber, C., Reiner, A. and Poole, A.R.: Isolation of dermatan sulfate proteoglycans from mature bovine articular cartilages. J Biol Chem 260 (1985) 6304-13.

Rosenblatt, S., Bassuk, J.A., Alpers, C.E., Sage, E.H., Timpl, R. and Preissner, K.T.: Differential modulation of cell adhesion by interaction between adhesive and counter-adhesive proteins: characterization of the binding of vitronectin to osteonectin (BM40, SPARC). Biochem J 324 (1997) 311-9.

Runyan, R.B. and Markwald, R.R.: Invasion of mesenchyme into three-dimensional collagen gels: a regional and temporal analysis of interaction in embryonic heart tissue. Dev Biol 95 (1983) 108-14.

Rupp, P.A.: Characterization of cirrin, a candidate gene for congenital heart defects in 3p- syndrome. Dissertation. Oregon Health and Sciences University, School of Medicine, Department of Molecular and Medical Genetics. Portland, OR, 1999.

*Rupp, P.A., *Fouad, G.T., Egelston, C.A., Reifsteck, C.A., Olson, S.B., Knosp, W.M., Glanville, R.W., Thornburg, K.L., Robinson, S.W. and Maslen, C.L.: Identification, genomic organization and mRNA expression of CRELD1, the founding member of a unique family of matricellular proteins. Gene 293 (2002) 47-57. * Co-first authors.

Santra, M., Mann, D.M., Mercer, E.W., Skorski, T., Calabretta, B. and Iozzo, R.V.: Ectopic expression of decorin protein core causes a generalized growth suppression in neoplastic cells of various histogenetic origin and requires endogenous p21, an inhibitor of cyclin-dependent kinases. J Clin Invest 100 (1997) 149-57.

Schmidt, G., Hausser, H. and Kresse, H.: Interaction of the small proteoglycan decorin with fibronectin. Involvement of the sequence NKISK of the core protein. Biochem J 280 (1991) 411-4.

Scholzen, T., Solursh, M., Suzuki, S., Reiter, R., Morgan, J.L., Buchberg, A.M., Siracusa, L.D. and Iozzo, R.V.: The murine decorin. Complete cDNA cloning, genomic organization, chromosomal assignment, and expression during organogenesis and tissue differentiation. J Biol Chem 269 (1994) 28270-81.

Schonherr, E., Levkau, B., Schaefer, L., Kresse, H. and Walsh, K.: Decorin-mediated signal transduction in endothelial cells. Involvement of Akt/protein kinase B in up-regulation of p21(WAF1/CIP1) but not p27(KIP1). J Biol Chem 276 (2001) 40687-92.

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

Sinning, A.R.: Partial purification of HLAMP-1 provides direct evidence for the multicomponent nature of the particulate matrix associated with cardiac mesenchyme formation. J Cell Biochem 66 (1997) 112-22.

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

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

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

Stankiewicz, P. and Lupski, J.R.: Genome architecture, rearrangements and genomic disorders. Trends Genet 18 (2002) 74-82.

Sutherland, H.F., Kim, U.J. and Scambler, P.J.: Cloning and comparative mapping of the DiGeorge syndrome critical region in the mouse. Genomics 52 (1998) 37-43.

Termine, J.D., Kleinman, H.K., Whitson, S.W., Conn, K.M., McGarvey, M.L. and Martin, G.R.: Osteonectin, a bone-specific protein linking mineral to collagen. Cell 26 (1981) 99-105.

Terrett, J.A., Newbury-Ecob, R., Cross, G.S., Fenton, I., Raeburn, J.A., Young, I.D. and Brook, J.D.: Holt-Oram syndrome is a genetically heterogeneous disease with one locus mapping to human chromosome 12q. Nat Genet 6 (1994) 401-4.
Tsuda, T., Philp, N., Zile, M.H. and Linask, K.K.: Left-right asymmetric localization of flectin in the extracellular matrix during heart looping. Dev Biol 173 (1996) 39-50.

Turbay, D., Wechsler, S.B., Blanchard, K.M. and Izumo, S.: Molecular cloning, chromosomal mapping, and characterization of the human cardiac-specific homeobox gene hCsx. Mol Med 2 (1996) 86-96.

Verjall, M.a.N., J.D.: A patient with a partial deletion of the short arm of chromosome 3. Am J Disease Children 132 (1978) 43-45.

Waldo, K., Miyagawa-Tomita, S., Kumiski, D. and Kirby, M.L.: Cardiac neural crest cells provide new insight into septation of the cardiac outflow tract: aortic sac to ventricular septal closure. Dev Biol 196 (1998) 129-44.

Winnemoller, M., Schmidt, G. and Kresse, H.: Influence of decorin on fibroblast adhesion to fibronectin. Eur J Cell Biol 54 (1991) 10-7.

Winnemoller, M., Schon, P., Vischer, P. and Kresse, H.: Interactions between thrombospondin and the small proteoglycan decorin: interference with cell attachment. Eur J Cell Biol 59 (1992) 47-55.

Wrana, J.L.: Regulation of Smad activity. Cell 100 (2000) 189-92.

Xue, Y., Gao, X., Lindsell, C.E., Norton, C.R., Chang, B., Hicks, C., Gendron-Maguire, M., Rand, E.B., Weinmaster, G. and Gridley, T.: Embryonic lethality and vascular defects in mice lacking the Notch ligand Jagged1. Hum Mol Genet 8 (1999) 723-30.

Yamagishi, H., Garg, V., Matsuoka, R., Thomas, T. and Srivastava, D.: A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects. Science 283 (1999) 1158-61.

Yamaguchi, Y., Mann, D.M. and Ruoslahti, E.: Negative regulation of transforming growth factor-beta by the proteoglycan decorin. Nature 346 (1990) 281-4.

Yamakawa, K., Huo, Y.K., Haendel, M.A., Hubert, R., Chen, X.N., Lyons, G.E., and Korenberg, J.R.: DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. Hum Mol Genet 7 (1998) 227-37.

Yost, J.C. and Sage, E.H.: Specific interaction of SPARC with endothelial cells is mediated through a carboxyl-terminal sequence containing a calcium-binding EF hand. J Biol Chem 268 (1993) 25790-6.

Young, B.A., Wang, P. and Goldblum, S.E.: The counteradhesive protein SPARC regulates an endothelial paracellular pathway through protein tyrosine phosphorylation. Biochem Biophys Res Commun 251 (1998) 320-7.

Yutzey, K.E., Rhee, J.T. and Bader, D.: Expression of the atrial-specific myosin heavy chain AMHC1 and the establishment of anteroposterior polarity in the developing chicken heart. Development 120 (1994) 871-83.

Appendix I

Vector Primers

T7	5'-TAATACGACTCACTATAGGG-3'
T7(c)	5'-TAATACGACTCACTATAGGGC-3' NOTE: for use only with Clontech vectors. Do not use with pMOS.
T3	5'-AATTAACCCTCACTAAAGGG-3'

Mouse Sequencing and PCR Primers

E1-2F	5'-TCTCCTCCTCCCCATCCTTC-3'
E1-2R	5'-GTGTTTCCACCCCCGAAG-3'
E2-3F	5'-TGGGAGGAAGAGAAGTTGTC-3'
E2-3F.2	5'-GGCTCCTCTAGTATTCCACTG-3'
E2-3F.3	5'-TATATATTAATTTCTTAAATTG-3'
E2-3F.4	5'-GTAGTGAGAACCTGATTACAACC-3'
E2-3F.5	5'-AAAACACAGAAGTGTGCCC-3'
E2-3R	5'-TCCACCAGGCGGGTCTCA-3'
M.Ex3-4F	5'-CTGGTGGAGGTGCTGGAG-3'
M.Ex3-4R	5'-TCGGGGGCTTCCTGCT-3'
M.Ex5-6F	5'-GCACTGTGACTGCCAAGC-3'
M.Ex5-6R	5'-CTGCAGACAATGGGATTCCT-3'
M.Ex8-9F	5'-CTGGGCTGTATGGGAGCAG-3'
Mexon1.seq	5'-CCGGCAGGTATGACACGGATG-3'
Mintron1.seq	5'-GAATGTGTCTAGCTCTGTAC-3'

Mintron2.1seq	5'-CAGTGGAATACTAGAGGAGCC-3'
Neo1R	5'-CAATTTAAGAAATTAATATATAAAC-3'
Neo1R2	5'-CAATTTAAGAAATTAATATATA-3'
Linker "T"	5'-GATCCACTAGTAGATCTGAGCT-3'
Linker "B"	5'-CAGATCTACTAGTG-3'
Genotype2.F	5'-CCAGTCAAAAACCACAGAGAGGG-3'

Yeast two-hybrid cloning and sequencing primers

Y2H3F	5'-TGTCATACCTGCCGGGG-3'
Y2H2R	5'-GCCCCCTCGTGTCCC-3'
5' AD-LD Amplimer	5'-CTATTCGATGATGAAGATACCCCACCAAACC-3'
3' AD-LD Amplimer	5'-GTGAACTTGCGGGGTTTTTCAGTATCTACGATT-3'

Appendix II

Table 1: Southern Blot matrix outlining various enzymes, probes and wash conditions used at the murine *Creld1* locus.

ENZYME(S)	PROBE	LABEL	WASH CONDITIONS	RESULTS
ClaI	1.8	^{32}P	moderate	non-specific
ClaI, SalI, XhoI, SacII	1.8	^{32}P	moderate	blank
ClaI, SalI, XhoI, SacII	1.8	^{32}P	moderate	non-specific
ClaI, SalI, XhoI, SacII	B	^{32}P	moderate	non-specific
ClaI, SalI, XhoI, SacII	B	^{32}P	moderate	blank
NdeI	B	ECL Nuc.	moderate/high	high background
NdeI	B	AP Direct	moderate/high	high background
NdeI, XhoI	B	AP Direct	moderate/high	non-specific
NdeI, XhoI	1.8	AP Direct	moderate/high	non-specific; +++ signal
NdeI, XhoI	N/S	AP Direct	moderate/high	very low signal
ClaI, SalI, XhoI, SacII	E1-2F/R	AP Direct	moderate/high	non-specific
NdeI, XhoI	B	AP Direct	high	non-specific
NdeI, XhoI,	N/S	AP Direct	high	non-specific
NdeI, XhoI, SacII, ClaI	M.Ex3-4F/R	^{32}P	moderate/high	Nde I positive
NdeI, XhoI, SacII, ClaI	M.Ex3-4F/R	^{32}P/preabsorbed total mouse genomic DNA	moderate/high	non-specific
BglII, HindIII, EcoRI, SpeI	M.Ex3-4F/R	^{32}P	moderate/high	Positive for BglII

LamdaFix II multiple cloning sites:

XbaI – Sac I – Not I – Sac I – Sal I –*Genomic Insert* – Sal I – Sac I – NotI – Sac I – Xba I
T7→ ←T3

Appendix II

Figure 1

Amino acid sequence comparison between human and mouse translation of *CRELD1* cDNA clones. *CRELD1* is very highly conserved, with 91% identity (94% similarity) at the amino acid level. WE domain is underlined.

ELSEVIER

Gene 293 (2002) 47–57

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Identification, genomic organization and mRNA expression of *CRELD1*, the founding member of a unique family of matricellular proteins

Paul A. Rupp[a,1], Gameil T. Fouad[a,1], Carley A. Egelston[a], Carol A. Reifsteck[a], Susan B. Olson[a], Wendy M. Knosp[a], Robert W. Glanville[b,d], Kent L. Thornburg[b,c], Susan W. Robinson[b], Cheryl L. Maslen[a,b,c,*]

[a]*Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, OR 97201, USA*
[b]*Heart Research Center, Oregon Health and Science University, Portland, OR 97201, USA*
[c]*Department of Medicine, L465, Oregon Health and Science University, 3181 S.W. Sam Jackson Park Road, Portland, OR 97201, USA*
[d]*Oregon Medical Laser Center, Providence St. Vincent Hospital,Portland OR 97225, USA*

Received 20 February 2002; received in revised form 1 May 2002; accepted 14 May 2002
Received by K. Gardiner

Abstract

We have isolated and characterized a unique gene that encodes a highly conserved membrane bound extracellular protein that defines a new epidermal growth factor-related gene family. The *CRELD1* (Cysteine-Rich with EGF-Like Domains 1) gene (previously known as cirrin) was cloned from a human chromosome 3 BAC. Mapping of the gene confirmed its position at chromosome 3p25.3. The gene is ubiquitously expressed in early development and later becomes more markedly expressed in the developing heart, limb buds, mandible and central nervous system. Expression persists in adulthood in most tissues. Sequence analysis suggests that this is a cell adhesion protein. The mouse orthologue was cloned and mapped to the syntenic region of mouse chromosome 6. Orthologues or homologues have also been identified for cow, Chinese hamster, *Drosophila* and *Caenorhabditis elegans*. The *CRELD1* gene is deleted in the human cytogenetic disorder 3p − syndrome and is in the region of loss of heterozygosity for several types of cancer. A potential role for this protein in these disorders is discussed.

Keywords: Cell adhesion; Chromosome 3p25; Cirrin, EGF superfamily; Extracellular; Transmembrane domain

1. Introduction

Epidermal growth factor (EGF)-like repeats are a class of cysteine rich domains that are found in a wide variety of extracellular proteins. These may be completely secreted into the extracellular space, or have transmembrane domains that tether the otherwise extracellular protein to the cell surface. EGF-like domains are often implicated in protein–protein interactions (Rao et al., 1995), but the functions of the various types of proteins are diverse and include extracellular matrix structural components, growth factors and related proteins, transmembrane receptors, cell adhesion molecules, and signaling proteins (Davis, 1990). Many of these molecules play significant roles in development (Engel, 1989). The extent of the EGF superfamily is not yet known and the addition of new families, such as the one described here, expands the range of molecules related through this common motif.

Here we present the cloning, sequencing, mapping and characterization of gene expression of a new member of the EGF superfamily designated *CRELD1* (Cysteine-Rich with EGF-Like Domains 1). Analysis of the protein sequence suggests that the *CRELD1* gene encodes a cell adhesion molecule. Mapping of *CRELD1* with respect to cytogenetic breakpoints in 3p − syndrome shows that it is frequently included in the 3p deletion. It is also included in a common region of loss of heterozygosity for several different types of cancer. In addition, we define the highly conserved nature of

Abbreviations: AVSD, atrioventricular septal defect; BAC, bacterial artificial chromosome; bp, base pair(s); cbEGF, calcium binding EGF-like domain; CHD, congenital heart defect; CNS, central nervous system; cR, centiRays; CRELD, cysteine rich with EGF-like domains; *CRELD1*, human CRELD1 gene; *Creld1*, mouse CRELD1 orthologue; ECM, extracellular matrix; EGF, epidermal growth factor; EST, expressed sequence tag; FISH, fluorescent in situ hybridization; kb, kilobase(s); nt, nucleotide(s); ORF, open reading frame; PCR, polymerase chain reaction; *pI*, isoelectric point; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl/0.015 M Na_3 citrate, pH 7.6; UTR, untranslated region

* Corresponding author. Tel.: +1-503-494-2011; fax: +1-503-494-6986.
E-mail address: maslenc@ohsu.edu (C.L. Maslen).

[1] These authors made equal contributions to the work presented here.

0378-1119/02/$ - see front matter
PII: S0378-1119(02)00696-0

this novel gene in the context of orthologues and homologues discovered in several diverse species ranging from *Homo sapiens* to *Caenorhabditis elegans*.

2. Materials and methods

2.1. cDNA sequence analysis

A screen of EST clones mapping to human chromosome 3p24.2–25 revealed a cDNA with sequence similarities to the fibrillins and other ECM proteins (Timmers et al., 1996). This cDNA clone was originally selected for further characterization as a candidate gene for the second Marfan syndrome locus in 3p25 (Collod et al., 1994). At the time the complete cDNA sequence was not available in the database so the corresponding cDNA clone was obtained from the Whitehead Institute (accession number WI11041) and sequenced in its entirety. The sequence was confirmed from normal human fibroblast cDNA.

2.2. Determination of genomic organization

We determined the genomic organization of *CRELD1* by PCR amplification and sequencing of genomic DNA using BAC 172I17 as the template (kindly provided by R. Moses and J. Henja). Intron–exon boundaries were subsequently confirmed by sequence analysis of genomic DNA from normal control subjects. Introns were completely sequenced with the exception of introns 2, 4 and 6, for which the intron–exon boundaries and several hundred internal bases were sequenced. Highly repetitive elements in the gene sequence were identified using the RepeatMasker program (http://repeatmasker.genome.washington.edu/cgi-bin/RepeatMasker).

2.3. Cloning of cow and mouse CRELD1 gene

BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST/) using human *CRELD1* sequences identified *Bos taurus* (cow) and mouse ESTs in the database, although complete cDNA sequences were not available. The cDNA sequences were determined by sequencing full-length cDNA clones for each species. The cow clone was obtained from the BACPAC resource, and the mouse clone was obtained from Research Genetics (I.M.A.G.E. clone 1547690). Genomic organization for mouse *Creld1* was determined by designing PCR primers to amplify across each of the presumptive intron–exon boundaries as defined in the human gene.

2.4. Homology and domain searches

Homology searches were done using the standard nucleic acid and protein BLAST algorithms (blastn and blastp) located at the National Center for Biotechnology Information (NCBI) homepage (http://www.ncbi.nlm.nih.gov/BLAST/). The blastp searches were performed with the translation of the sequence from the complete open reading frame and also as individual exons or domains using the basic program parameters. Protein topology predictions were done using the Simple Modular Architecture Research Tool (SMART) program (http://smart.embl-heidelberg.de/). The amino acid sequence was also analysed through the TMHMM algorithm to confirm the prediction of transmembrane domains and predict membrane topology (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Optimal alignments of human, cow, mouse, *Cricetulus griseus* (Chinese hamster), *Drosophila melanogaster* and *C. elegans* sequences were done using the multiple sequence alignment program CLUSTAL W 1.81 (http://www.ch.embnet.org/software/ClustalW.html), and identity and homology annotations were done in Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Potential sites of phosphorylation were identified using NetPhos 2.0 (http://www.cbs.dtu.dk/services/NetPhos/).

2.5. Radiation hybrid mapping

The GeneBridge 4 Radiation Hybrid Panel (Research Genetics) was queried to confirm the chromosomal location of human *CRELD1*. The chromosomal location of the mouse orthologue was determined using a mouse/hamster whole genome radiation hybrid mapping panel obtained from Research Genetics.

2.6. Northern blot hybridization

Human tissue Northern blots (Clontech) were probed with a cDNA fragment encompassing the entire coding sequence of *CRELD1* including 321 bp of 5′-UTR and 357 bases of 3′-UTR. The probe was labeled using the Gene Images labeling module (Amersham). Hybridization was carried out at 65 °C for 12 h. Blots were washed at a final stringency of 65 °C for 15 min in 0.1× SSC/0.1% SDS. *CRELD1* transcripts were detected using the Gene Images CDP-Star detection module (Amersham). The blots were stripped and then probed under the same conditions with a β-actin probe to check for equality of sample loading. Blots were exposed to Hyperfilm ECL (Amersham) for visualization of the signal.

2.7. Whole-mount in situ hybridizations

Chick embryos were harvested at 66, 72, and 90 h time points. Extraembryonic membranes were removed and embryos fixed in 4% paraformaldehyde at room temperature before dehydration with methanol for storage at 4 °C. Embryos were rehydrated with PBS before treatment with proteinase K and subsequent hybridization with probe. Antisense and sense RNA probes containing digoxigenin-11-UTP were created by transcriptional reactions from a pGEM-4Z construct containing nucleotides (−)321 to 731 of human *CRELD1*. The antisense probe was shown to cross-react with chick mRNA by Northern blot analysis

(not shown). Prior to transcription of probes with T7 or SP6 polymerases, the vector was linearized with either *Bam*HI or *Acc*I restriction enzymes to give distinct length RNA transcripts. Hybridization with 500 ng/ml of probe was carried out at 70 °C for a minimum of 17 h. Single-stranded RNA was removed with RNase A treatment before hybridization with alkaline phosphatase conjugated anti-digoxigenin antibody. Color reactions were quenched with numerous washes of PBS. Stained embryos were fixed with 4% paraformaldehyde prior to photographing and staging.

2.8. Fluorescent in situ hybridization (FISH) chromosome analysis

We used FISH analysis to determine if the *CRELD1* gene was deleted in 3p − syndrome, and to localize the gene relative to the critical region for congenital heart defects as defined by Green et al. (2000). The cell lines used for this study were obtained from the NIGMS Human Mutant Cell Repository (Coriell Institute). Four cell lines derived from unrelated patients with 3p − syndrome were obtained from the repository. Two of the cell lines, designated GM10922 and GM07873, are from patients who had an associated congenital heart defect (AVSD), and one cell line, GM10985, is from a patient who did not have a heart defect. GM10922 and GM10985 (Ramer et al., 1989) were previously used in the molecular genetic dissection of the critical region for CHD associated with 3p − syndrome (Phipps et al., 1994; Drumheller et al., 1996; Green et al., 2000). The fourth cell line, GM11403, was derived from a stillborn infant with 3p − syndrome, but there was no information available regarding a possible heart defect. Metaphase spreads from each cell line were hybridized with an 11 kb *CRELD1* genomic DNA probe labeled with digoxigenin. Since the *CRELD1* probe target was small, localization was confirmed using two different colors for detection (rhodamine, red and FITC, yellow) in separate FISH experiments.

3. Results

3.1. Characterization of human CRELD1 cDNA

CRELD1 cDNA is 2077 bp in length, with an open reading frame of 1263 nt (Fig. 1). At the proposed translation-initiating methionine there is a Kozak consensus sequence with a 'strong context' for being the translational start site (Kozak, 1996). The presence of several upstream in-frame stop codons is consistent with the assignment of the ATG initiator codon. Additional sequences upstream are indicative of a eukaryotic translational start domain (Ganoza and Louis, 1994). The 5′- and 3′-UTRs are 365 and 449 bp, respectively. An in-frame termination codon (TAA) begins at nucleotide 1261 and a consensus polyadenylation site (AATAAA) starts at nucleotide 1689. This sequence data has been deposited with GenBank under accession number AF452623.

3.2. Genomic organization

The intron–exon structure of the *CRELD1* gene was determined by DNA sequence analysis of PCR amplified fragments from BAC 172I17 and verified by sequence analysis of PCR amplified genomic DNA. There are 11 confirmed coding exons encompassed by approximately 12 kb of genomic DNA. 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. There are repetitive elements (SINE and LINE class) in introns 1, 2, 3, 4 and 9, and a simple hexanucleotide repeat $(gggaga)_n$ in intron 6. The major 2.1 kb transcript utilizes ten exons (exons 1–10). The positions of the intron–exon boundaries for this transcript are shown in Fig. 1. There is an alternatively spliced exon (exon 9b) embedded in intron 9, which is not included in the major transcript. The genomic organization is shown in Fig. 2.

3.3. Protein sequence analysis

The major transcript of the *CRELD1* gene encodes a 420-amino-acid protein. CRELD1 is predicted to be an acidic protein with a theoretical p*I* of 4.77. Sequence analysis using the SMART algorithm shows that there are several overlapping motifs, which complicates the picture of protein domain structure. Five alternative domain organizations are predicted depending on the positioning of boundaries between domains. The models differ from each other in that they contain either tandem arrays of EGF and/or calcium binding EGF domains, or EGF/calcium binding EGF domains and furin cysteine-rich domains. For the sake of clarity only one model is presented (Fig. 3). The protein has a unique region that is rich in glutamic acid and tryptophan residues that we have named the WE domain (W = tryptophan, E = glutamic acid). There is a poorly conserved low complexity proline-rich region at the beginning of this domain.

Computer analyses predict a secretion signal sequence at the amino-terminus with a cleavage site between amino acid residues 29–30, and two tandem type III transmembrane domains between amino acid residues 363–382 and 387–406. The deduced amino acid sequence with annotated sequence elements is shown in Fig. 1. The presence of an amino-terminal secretion signal and carboxyl-terminal type III transmembrane domains indicates a complex membrane topology (von Heijne and Gavel, 1988). This model suggests that amino acids 1–361 and 407–420 reside in the extracellular space, held in place by the two transmembrane helices with a short intervening cytoplasmic loop. A diagrammatic representation of the protein structure and topology is shown in Fig. 3.

In addition to recognizable protein domains, other

sequence-based elements are indicated. There are two consensus sites for *N*-linked glycosylation (NXS/T), one in the WE domain and one in the first EGF domain. Several high probability phosphorylation sites for serine, threonine and tyrosine residues were also identified. Positions of predicted phosphorylation sites with probability scores > 0.90 are indicated in Fig. 1. All of the potentially phosphorylated residues are in extracellular domains.

3.4. CRELD1 is expressed in multiple tissues

Human fetal and adult multiple tissue Northern blots were hybridized with a *CRELD1*-specific cDNA probe (Fig. 4A,B). A prominent transcript of 2.1 kb was observed in all tissues analysed. The highest levels of expression were in fetal lung, liver, kidney, adult heart, brain and skeletal muscle. Lower levels of expression were observed in

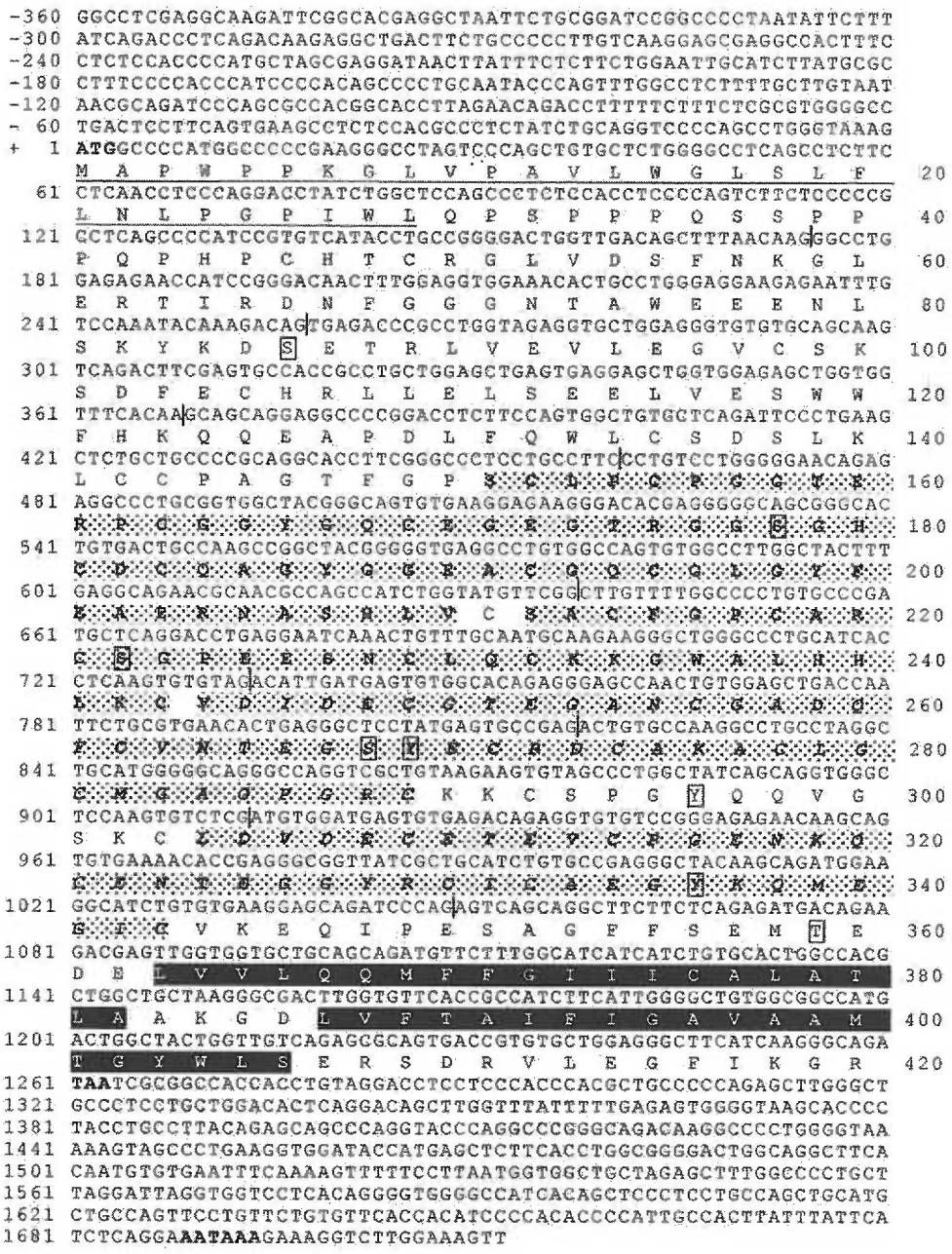


Fig. 1. The complete cDNA sequence and amino acid translation for the *CRELD1* gene. Nucleotide numbering is along the left border, numbering for the amino acid sequence is along the right border. The single-letter amino acid code is under the first nucleotide of the corresponding codon. For the DNA sequence, the initiation and termination codons, and the polyadenylation signal sequences are in bold typeface. Vertical lines mark intron–exon boundaries. For the amino acid sequence, the cleavable secretion signal sequence (amino acids 1–29) is underlined. The four predicted EGF-like domain sequences are shaded, with the two calcium binding EGF-like domains italicized. The two predicted type III transmembrane domains are in dark shading with reverse typeface. Amino acid residues that are likely to be phosphorylated (>90% probability) are in boxes.

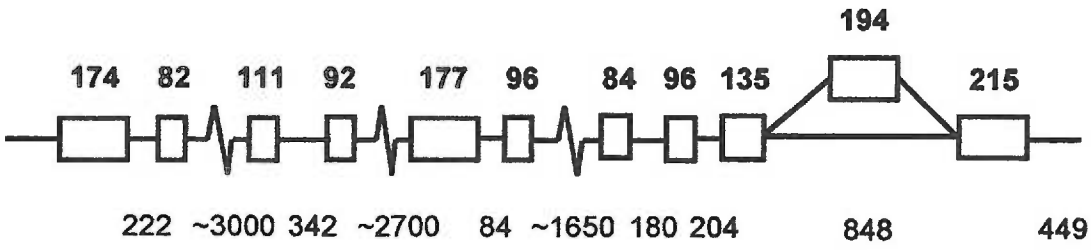


Fig. 2. Diagrammatic representation of the intron–exon boundary structure of *CRELD1*. The solid lines represent introns, the open boxes are exons. The sizes, in base pairs, are below the corresponding introns and above the corresponding exons.

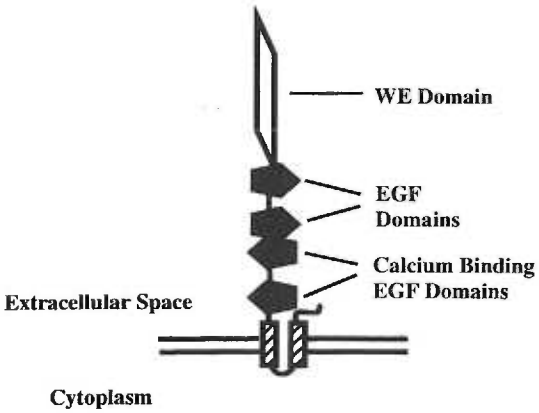


Fig. 3. Diagrammatic representation of CRELD1 showing the predicted domain structure and cell membrane topology. The two type III transmembrane domains are shown as hatched rectangles. The extracellular domains are labeled.

placenta, fetal brain, and adult lung, liver, kidney and pancreas. The 2.1 kb transcript is consistent in size with the cDNA clone characterized here, confirming that it represents the full-length cDNA of the most prominently expressed *CRELD1* variant. The same probe was used on a multiple tissue Northern array blot (Clontech RNA Master Blot) with 50 tissues represented. A signal was detected for all tissues included in the array (not shown). The most prominent signals were from caudate nucleus, cerebellum, cerebral cortex, frontal lobe, pituitary gland, adrenal gland, thyroid gland and kidney. In addition to the tissues analysed here, *CRELD1* is represented in a wide variety of cDNA libraries including those derived from eye, tonsil, blood, breast, skin, esophagus, colon and cervix, as identified from ESTs containing the *CRELD1* sequence. These data indicate that most fetal and adult tissues express *CRELD1*.

To get an overall view of the embryonic expression pattern of *CRELD1* we performed whole-mount in situ hybridizations on developing chick embryos. *CRELD1* expression was evident at all time points. Localized high levels of expression were seen in the developing heart, limb buds, mandible, branchial arches and brain (Fig. 5).

3.5. Transcription of multiple splice variants

There is considerable evidence that there are several

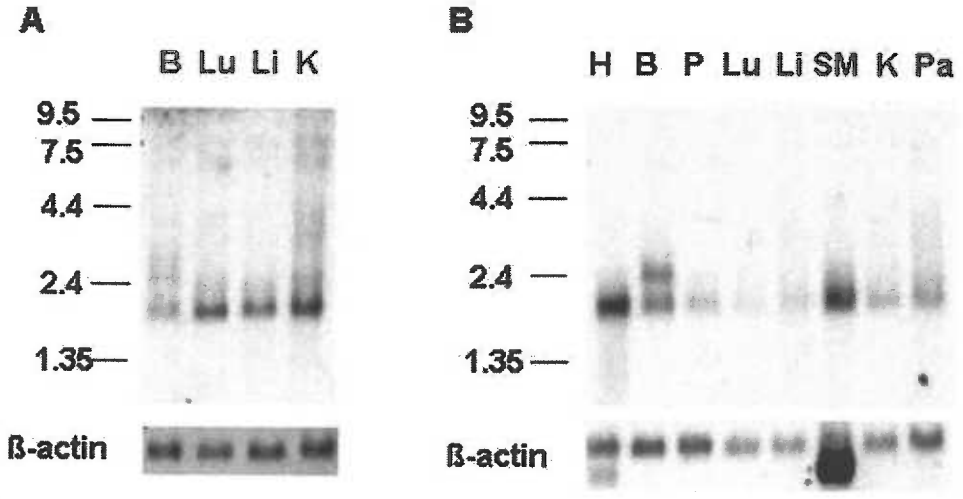


Fig. 4. Northern blot analysis of poly(A)$^+$ RNA from human fetal and adult tissues. (A) Fetal tissue Northern blot hybridized with a *CRELD1*-specific cDNA probe. Each lane contains 2 μg mRNA isolated from B, fetal brain; Lu, fetal lung; Li, fetal liver; K, fetal kidney. (B) Adult tissue Northern blot, also hybridized with the *CRELD1* probe. Each lane contains 2 μg mRNA isolated from H, heart; B, brain; P, placenta; Lu, lung; Li, liver; SM, skeletal muscle; K, kidney; Pa, pancreas. Note the 2.1 kb transcript in all lanes. Adult brain expresses a prominent second transcript of 2.5 kb. Under each blot is the same blot hybridized with a human β-actin cDNA probe used as a control to assess the relative amounts of RNA present in each lane.

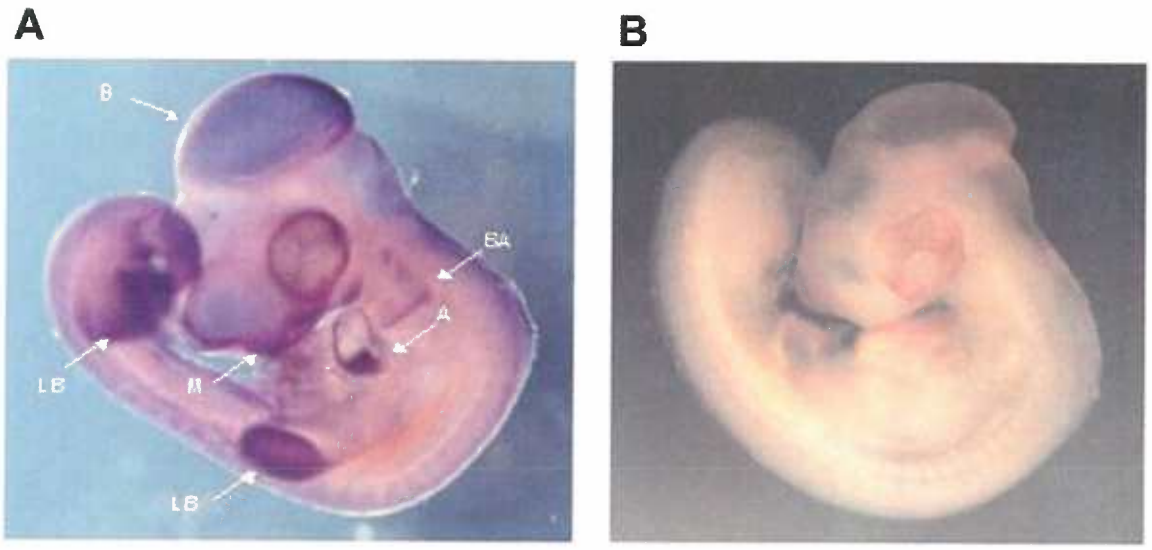


Fig. 5. Whole-mount in situ hybridizations of chick embryos stained for *CRELD1*gene expression. (A) Whole chick embryo after in situ hybridization with a *CRELD1*-specific antisense probe. Left view, HH stage 22, magnification 12.5×. Note staining in the developing brain (B), limb buds (LB), mandible (M), cardiac atrial muscle and cushion tissue (A), and branchial arches (BA). (B) Chick embryo after in situ hybridization with the sense strand from the *CRELD1*-specific probe as a negative control. Left view, HH stage 22, magnification 12×. Note the absence of any staining.

alternatively spliced transcripts produced from *CRELD1*. Northern blot analysis of poly(A)$^+$ RNA from adult brain shows an alternative transcript of 2.5 kb in equal abundance to the 2.1 kb transcript. The 2.5 kb transcript was not detected by the same probe in any other tissue examined including fetal brain, which primarily expresses the 2.1 kb transcript. There is another alternative transcript of about 2.3 kb produced in regions of the mature heart (Fig. 6). This alternate transcript was not detected in total adult or fetal heart mRNA, but became apparent on examination of subsections of mature heart including right ventricle, left ventricle, right atrium, left atrium, apex and aorta. Additional splice variants were also suggested by the presence of less abundant larger transcripts faintly visible on the fetal multiple tissue Northern blot. The constitution of these alternate transcripts has not been determined.

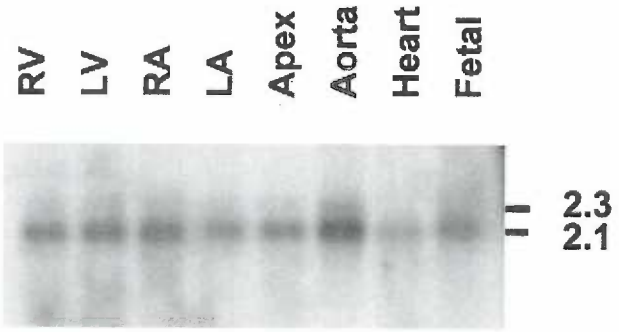


Fig. 6. Northern blot analysis of poly(A)$^+$ RNA isolated from specific regions of human heart, using *CRELD1* cDNA as the probe. Each lane contains 2 μg mRNA isolated from RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; apex; aorta; heart, total adult heart; fetal, total fetal heart. There are two transcripts (2.1 and 2.3 kb) visualized for all cardiovascular regions represented on the blot, although the less abundant 2.3 kb transcript is not visible in the adult and fetal whole heart mRNAs.

Sequence analysis of genomic DNA identified an open reading frame of 194 bp embedded in intron 9. Further investigation indicated that this ORF is an alternatively spliced exon, which we will refer to as exon 9b. Utilization of exon 9b alters the reading frame and introduces a new 73 amino acid sequence at the carboxyl-terminus of the predicted protein that replaces the 71 amino acids encoded by exon 10. The first eight codons of exon 10 are included in the mRNA, with the ninth codon as the termination codon in the new reading frame. Inclusion of exon 9b results in a protein that has significantly different characteristics, as there are no predicted transmembrane domains in the altered amino acid sequence. The unique carboxyl-terminus is not similar to any other sequence in the public domain databases and there are no recognizable protein motifs. Northern array blot analysis using an exon 9b-specific probe shows that transcripts utilizing this alternate exon are expressed in brain, jejunum, ileum, ilocecum, appendix, thymus, lymph node, fetal spleen, fetal thymus and fetal lung (data not shown). Inclusion of exon 9b has been ruled out as the basis for the 2.3 kb transcript expressed in heart.

3.6. Human CRELD1 maps to chromosome 3p25.3 and is deleted in 3p − syndrome

Query of the GeneBridge 4 Radiation Hybrid panel placed *CRELD1* on the Whitehead framework map proximal to *AFMA216ZG1* and distal to *D3S1263* on chromosome 3p25. This is consistent with the map location determined for I.M.A.G.E. consortium clone 137340, which corresponds to EST WI-8719. Sequence analysis

confirmed that clone 137340 contains the *CRELD1* gene. Assembly of a BAC contig map of 3p25.3 placed WI-8719 between markers *D3S1317* and *D3S1597* (Hejna et al., 2000).

Comparison of the map location of *CRELD1* with the detailed map of the congenital heart disease critical region ascertained by breakpoint analysis of 3p − syndrome (Green et al., 2000) indicates that the *CRELD1* locus is distal to this region. A gap in the physical map leads to uncertainty as to the physical distance between these loci. FISH analysis demonstrated that the *CRELD1* gene is deleted in all four of the 3p − syndrome cell lines examined (not shown). Deletion of *CRELD1* in cell lines derived from 3p − patients with no congenital heart disease confirms that *CRELD1* lies outside the critical region for congenital heart defects, assuming complete penetrance of that trait. However, it should be noted that incomplete penetrance has been previously described for AVSD (Pierpont et al., 2000), and if this is true for AVSD associated with 3p − syndrome the critical region would extend telomerically and could include *CRELD1*.

Deletion of *CRELD1* in GM10985 refines the breakpoint region for that cell line. Phipps and co-workers had previously mapped the breakpoint to the region between *D3S601* and *D3S18* (Phipps et al., 1994). Our analyses show that the breakpoint must lie in the region between *D3S601* and *D3S1597*. The breakpoint analyses and mapping data are summarized in Fig. 7.

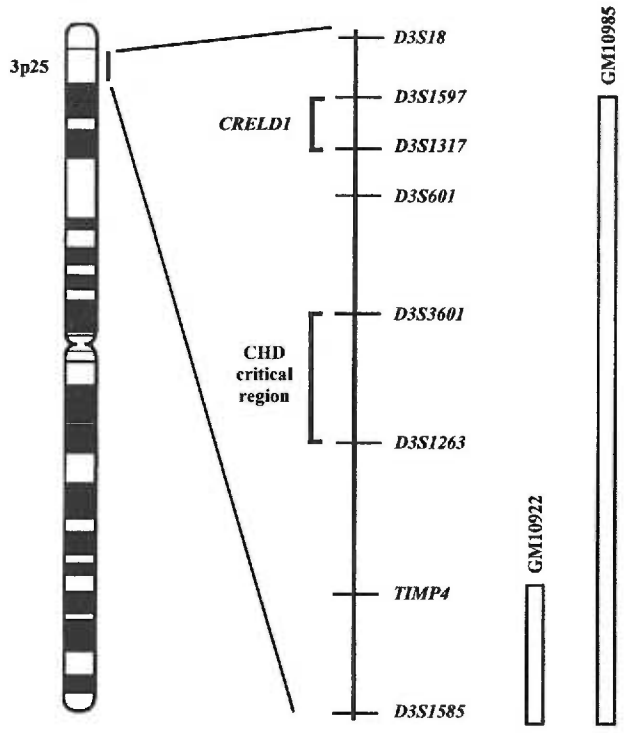


Fig. 7. Chromosomal localization of *CRELD1* with respect to a detailed map of markers in the 3p25 region and two known 3p − syndrome breakpoints. An ideogram of chromosome 3 is on the far left. The map location of the *CRELD1* gene and the congenital heart disease (CHD) critical region are shown relative to flanking markers. The open bars on the right show the positions of the chromosome terminus for GM10922 and GM10985 with respect to markers within 3p25. Since the exact positions of the deleted chromosome termini are unclear the most conservative interpretation is presented with the terminus extending to the most distal of the markers flanking the breakpoint region.

3.7. Identification of CRELD2

BLAST searches with the deduced amino acid sequence for CRELD1 identified a closely related gene on chromosome 22p13 (GenBank accession number NP_077300). Optimal alignment of the nucleic acid sequences for the coding regions of the two genes shows that they are 51% homologous. At the amino acid level they are 38% identical and 51% similar. The degree of sequence identity and similar organization of structural motifs indicates that these are homologous genes that encode distinct but closely related proteins. In particular, they share the unique WE domain, which has not been found in any non-homologous genes. In fact, the WE domains are the region of greatest similarity between the two proteins, where they are 55% identical and 75% similar in amino acid sequence. This indicates that CRELD1 is the founding member of a new protein family that has one other known member, designated CRELD2. The major distinction between the two proteins is that CRELD1 has two type III transmembrane domains at the carboxyl-terminus, whereas CRELD2 does not have any predicted transmembrane domains. This would suggest a significant functional difference between the two proteins with CRELD1 bound to the cell surface, but CRELD2 secreted freely into the extracellular space.

3.8. Cloning and characterization of the cow and mouse CRELD1 orthologues

The complete cDNA sequences of the cow and mouse orthologues of *CRELD1* were determined by sequencing of the relevant EST clones. The mouse orthologue has been designated *Creld1*. The open reading frames for the cow and mouse cDNAs are 89 and 88% identical to the human cDNA, respectively. Subsequent to this analysis sequence for the mouse *Creld1* cDNA became available as an anonymous sequence in the GenBank database (accession number BC010804), confirming our sequence results. Like the human gene, the cow and mouse genes both encode 420 amino acid proteins, which are 94 and 92% identical to human CRELD1, respectively, indicating that these are orthologues. Our analysis of the genomic organization of the mouse gene shows that it is comparable to that of human *CRELD1*, with ten coding exons of similar size and position to the ten major exons in the human gene.

3.9. Localization of mouse Creld1 to distal chromosome 6

Screening of a mouse/hamster radiation hybrid panel localized the mouse *Creld1* gene to a 26.6 cR interval between *D6Mit10* and *D6Mit329*, with highest anchor

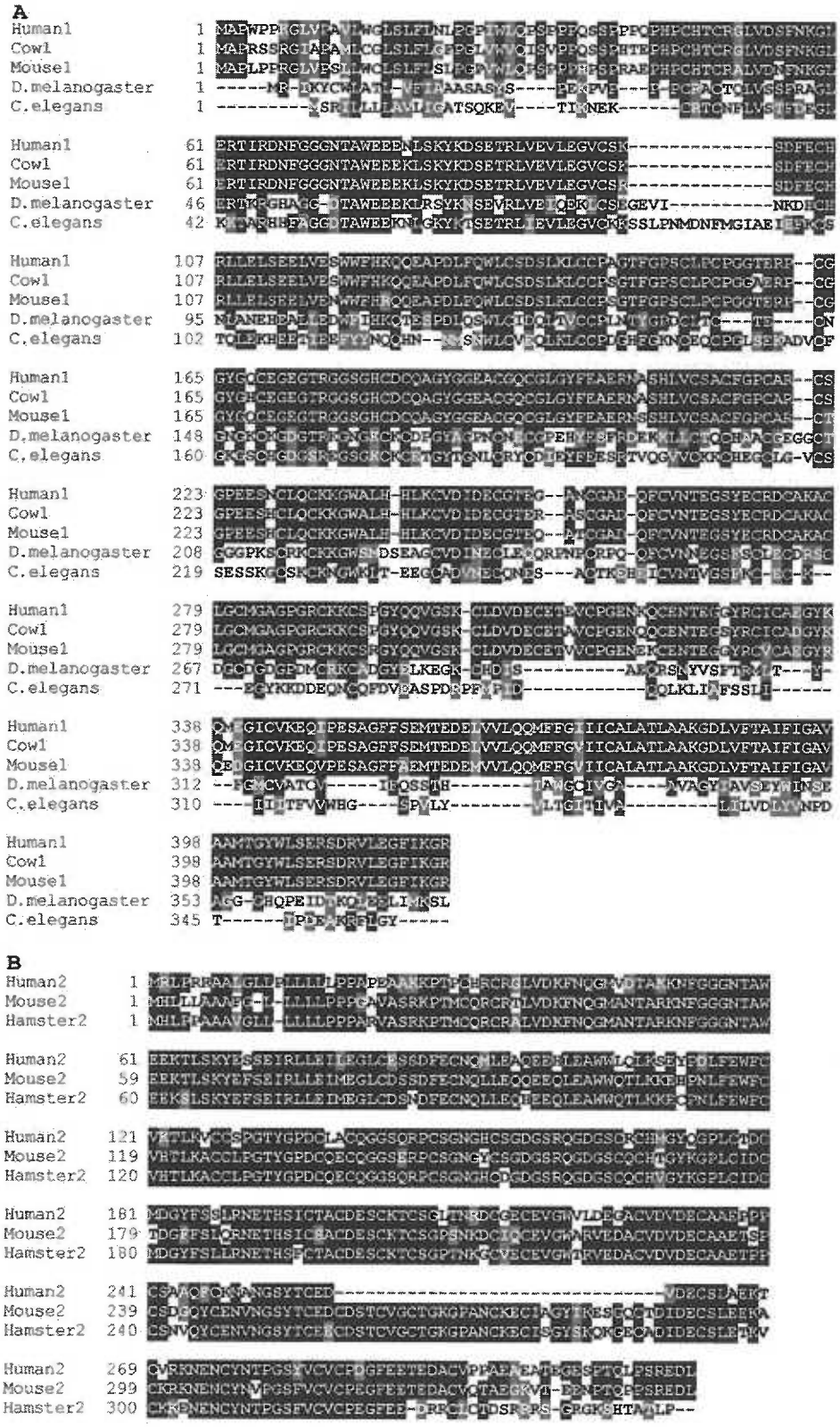


Fig. 8. Optimal alignments of the known *CRELD* homologues. Amino acid residues that are identical between species are dark shaded with reverse typeface. Conservative amino acid substitutions between species are light shaded with reverse typeface. Note the high degree of identity throughout both proteins. (A) Alignment of the *CRELD1* human, cow and mouse orthologues along with the *Drosophila* and *Caenorhabditis elegans* homologues. (B) Alignment of the human, mouse and *Cricetulus griseus CRELD2* orthologues.

LOD score of 12.5 to *D6Mit252*, indicating the gene is on distal chromosome 6. This finding is in agreement with predictions made from the human mapping data, which places human *CRELD1* on chromosome 3p25, an area known to be syntenic with distal mouse chromosome 6. The position of *Creld1* relative to flanking genes and markers is consistent with the order of loci flanking human *CRELD1*. Mapping of mouse *Creld1* to a region of synteny with the human *CRELD1* locus further supports the conclusion that the two genes are orthologues.

3.10. Conservation of CRELD genes across species

Cloning and characterization of the cow and mouse orthologues, and homology searches of gene and protein sequence databases show that *CRELD1* is highly conserved across species. Database searches have identified additional *CRELD* homologues in Chinese hamster, *Drosophila* and *C. elegans* genomes. There is 44% amino acid identity between CRELD1 and the product of gene *F09E8* from *C. elegans* (GenBank accession number Z73896). The *D. melanogaster* gene designated *CG11377* also encodes a protein that is similar to human CRELD1 (44% amino acid identity). The Chinese hamster *CRELD* homologue was originally characterized as the gene for a putative extracellular protein named HT protein (GenBank accession number U48852). This gene is most closely related to *CRELD2* with 67% amino acid identity, but only 47% amino acid identity with CRELD1. There is also a murine *Creld2* orthologue (GenBank accession number AK017880) that has 69% identity with human *CRELD2* and 83% identity with Chinese hamster HT protein/*CRELD2* homologue. Optimal alignments of the amino acid sequences for the human, cow, mouse, Chinese hamster, *Drosophila* and *C. elegans* *CRELD1* and *CRELD2* homologues are shown in Fig. 8. It should be noted that there is extensive identity between species throughout the molecule; the similarities are not restricted to the genetically mobile EGF domains. Of particular interest is the highly conserved unique WE domain. This domain has not been detected in any previously described proteins, and its function is unknown. The absolute conservation of a GG(N/D)TAWEE(E/K) nonapeptide in all WE domains characterized to date indicates that it may be a region important for interactions with other proteins.

4. Discussion

We describe here the identification and characterization of the founding member of a previously unknown protein family, which belongs to the EGF superfamily. This gene, previously known as cirrin (Maslen et al., 1999; Green et al., 2000; Pierpont et al., 2000), has been designated *CRELD1* by the Human Genome Nomenclature committee. Although the specific function of CRELD1 protein is elusive the predicted protein structure indicates that it is likely to be a cell adhesion molecule. Cell surface EGF-like proteins generally participate in cell-cell interactions, where they play roles in influencing growth and development (Engel, 1989). Hynes and Zhao identified the *Drosophila* *CRELD1* homologue (CT31762) as belonging to this class of proteins in a search of the *Drosophila* genome for proteins with domains prevalent in cell adhesion molecules (Hynes and Zhao, 2000). In their analysis the presence of calcium binding EGF-like domains served as an indicator of protein function.

In contrast to the EGF-like domains, the WE domain is unique to the CRELD protein family. It is virtually identical in the human, cow, mouse and Chinese hamster orthologues, and is highly conserved between other species with a surprising degree of identity between human CRELD1 and the *Drosophila* and *C. elegans* homologues (51% and 53% respectively). From the data presented here we conclude that the CRELD proteins constitute a new family of matricellular proteins.

Protein sequence analyses indicate that the CRELD proteins have the potential to be heavily phosphorylated. It is of note that the Chinese hamster *CRELD2* orthologue (HT protein) was originally cloned by screening of an expression library with an anti-phosphotyrosine antibody (GenBank reference and H. Chen, Allelix Biopharmaceuticals, personal communication). This supports the supposition that CRELD proteins are indeed subject to phosphorylation, at least on tyrosine residues. It is interesting that all of the potential phosphorylation sites, including those of low probability, are predicted to be extracellular. This suggests that CRELD phosphorylation is mediated by ecto-protein kinases (Redegeld et al., 1999). Phosphorylation of the extracellular domains of CRELD1 may be important to the adhesive responses of associated cells. Phosphorylation of soluble extracellular ligands, such as CRELD2, may alter receptor affinity or bioactivity.

Searches of the complete genomes of *C. elegans* and *Drosophila* show that they each contain a single CRELD homologue. In both cases the homologues are more closely related in amino acid sequence to human CRELD1 than CRELD2, although the *Drosophila* sequence does not have any predicted transmembrane domains at the carboxyl-terminus. The *C. elegans* sequence has a single predicted carboxyl-terminal transmembrane domain. Inasmuch as the *C. elegans* and *Drosophila* genomes have been completely sequenced these data indicate that the presence of two *CRELD* homologues in a genome is an evolutionary advance, with lower species maintaining a single locus.

A review of the literature demonstrates that chromosome 3p25 is a region of intense genetic activity. Numerous undefined genetic loci for a diverse range of disorders and traits have been mapped to this area, and there is considerable morbidity associated with this region. This includes multiple cancer-related genes (Deng et al., 1998; Maestro et al., 2000; Wistuba et al., 2000; Barghorn et al., 2001; Kayahara et al., 2001; Maitra et al., 2001; Martinez et al., 2001; Sekine et al., 2001; Simsir et al., 2001; Tschentscher et al., 2001), a second locus for Marfan syndrome (Collod et al., 1994), the

cerebrovascular condition known as Moyamoya disease (Ikeda et al., 1999), congenital cataracts (Hampe et al., 2001) and inflammatory bowel disease (Pras et al., 2001). In the course of this study we investigated *CRELD1* as a candidate gene for Marfan syndrome and were able to exclude it from that locus (unpublished data). However, the characteristics of the protein and location of the gene make *CRELD1* a compelling candidate for many of these disorders. Of particular interest is the close proximity of *CRELD1* to recently described loci for familial ovarian cancer (Sekine et al., 2001) and breast carcinoma (Maitra et al., 2001). *CRELD1* also resides in the defined intervals of loss of heterozygosity for nasopharyngeal carcinoma (Deng et al., 1998), lung cancer (Wistuba et al., 2000), and possibly other cancers with loss of heterozygosity in 3p25. Determination of involvement of *CRELD1* in allelic loss associated with malignant processes and characterization of CRELD1 as a cell adhesion molecule will be critical to understanding its potential role in cancer. Targeted deletion of *CRELD1* in a heterozygous knockout mouse will provide information as to the role deletion of *CRELD1* plays in 3p − syndrome and in normal development, as well as the potential for involvement in other diseases.

Acknowledgements

We thank Andrew Mendenhall, Darcie Babcock and Stephanie McInerney for technical assistance, Markus Grompe and Cynthia Timmers for access to sequence data and clone WI11041, Robb Moses and Jim Hejna for BAC clone 172I17 and related information. G.T.F. is a Predoctoral Fellow of the American Heart Association, Northwest Affiliate Inc. C.A.E. was a fellow of the M.J. Murdock Charitable Trust. This work has been supported by grants to C.L.M. from the American Heart Association, Northwest Affiliate Inc., the Medical Research Foundation of Oregon and the Gerlinger Foundation.

References

Barghorn, A., Komminoth, P., Bachmann, D., Rutimann, K., Saremaslani, P., Muletta-Feurer, S., Perren, A., Roth, J., Heitz, P.U., Speel, E.J., 2001. Deletion at 3p25.3–p23 is frequently encountered in endocrine pancreatic tumours and is associated with metastatic progression. J. Pathol. 194, 451–458.

Collod, G., Babron, M.C., Jondeau, G., Coulon, M., Weissenbach, J., Dubourg, O., Bourdarias, J.P., Bonaitipellie, C., Junien, C., Boileau, C., 1994. A second locus for Marfan syndrome maps to chromosome 3p24.2–p25. Nat. Genet. 8, 264–268.

Davis, C.G., 1990. The many faces of epidermal growth factor repeats. New Biol. 2, 410–419.

Deng, L., Jing, N., Tan, G., Zhou, M., Zhan, F., Xie, Y., Cao, L., Li, G., 1998. A common region of allelic loss on chromosome region 3p25.3–26.3 in nasopharyngeal carcinoma. Genes Chromosomes Cancer 23, 21–25.

Drumheller, T., McGillivray, C., Behrner, D., MacLeod, P., McFadden, D.E., Roberson, J., Venditti, C., Chorney, D., Chorney, M., Smith, D.I., 1996. Precise localisation of 3p25 breakpoints in four patients with the 3p − syndrome. J. Med. Genet. 33, 842–847.

Engel, J., 1989. EGF-like domains in extracellular matrix proteins: Localized signals for growth and differentiation? FEBS Lett. 251, 1–7.

Ganoza, M.C., Louis, B.G., 1994. Potential secondary structure at the translational start domain of eukaryotic and prokaryotic mRNAs. Biochimie 76, 428–439.

Green, E.K., Priestley, M.D., Waters, J., Maliszewska, C., Latif, F., Maher, E.R., 2000. Detailed mapping of a congenital heart disease gene in chromosome 3p25. J. Med. Genet. 37, 581–587.

Hampe, J., Lynch, N.J., Daniels, S., Bridger, S., Macpherson, A.J., Stokkers, P., Forbes, A., Lennard-Jones, J.E., Mathew, C.G., Curran, M.E., Schreiber, S., 2001. Fine mapping of the chromosome 3p susceptibility locus in inflammatory bowel disease. Gut 48, 191–197.

Hejna, J.A., Timmers, C.D., Reifsteck, C., Bruun, D.A., Lucas, L.W., Jakobs, P.M., Toth-Fejfel, S., Unsworth, N., Clemens, S.L., Garcia, D.K., Naylor, S.L., Thayer, M.J., Olson, S.B., Grompe, M., Moses, R.E., 2000. Localization of the Fanconi anemia complementation group D gene to a 200-kb region of chromosome 3p25.3. Am. J. Hum. Genet. 66, 1540–1551.

Hynes, R.O., Zhao, Q., 2000. The evolution of cell adhesion. J. Cell Biol. 150, F89–F96.

Ikeda, H., Sasaki, T., Yoshimoto, T., Fukui, M., Arinami, T., 1999. Mapping of a familial Moyamoya disease gene to chromosome 3p24.2–p26. Am. J. Hum. Genet. 64, 533–537.

Kayahara, H., Yamagata, H., Tanioka, H., Miki, T., Hamakawa, H., 2001. Frequent loss of heterozygosity at 3p25–p26 is associated with invasive oral squamous cell carcinoma. J. Hum. Genet. 46, 335–341.

Kozak, M., 1996. Interpreting cDNA sequences: Some insights from studies on translation. Mamm. Genome 7, 563–574.

Maestro, M.L., del Barco, V., Sanz-Casla, M.T., Moreno, J., Adrover, E., Izquierdo, L., Zanna, I., Fernandez, C., Redondo, E., Blanco, J., Resel, L., 2000. Loss of heterozygosity on the short arm of chromosome 3 in renal cancer. Oncology 59, 126–130.

Maitra, A., Wistuba, I.I., Washington, C., Virmani, A.K., Ashfaq, R., Milchgrub, S., Gazdar, A.F., Minna, J.D., 2001. High-resolution chromosome 3p allelotyping of breast carcinomas and precursor lesions demonstrates frequent loss of heterozygosity and a discontinuous pattern of allele loss. Am. J. Pathol. 159, 119–130.

Martinez, A., Walker, R.A., Shaw, J.A., Dearing, S.J., Maher, E.R., Latif, F., 2001. Chromosome 3p allele loss in early invasive breast cancer: Detailed mapping and association with clinicopathological features. Mol. Pathol. 54, 300–306.

Maslen, C.L., Rupp, P.A., Olson, S.B., Reifsteck, C.A., Thornburg, K.L., Glanville, R.W., 1999. Characterization of cirrin, a new extracellular protein gene that is associated with endocardial cushion defects in 3p − syndrome. Am. J. Hum. Genet. 65, A43.

Phipps, M.E., Latif, F., Prowse, A., Payne, S.J., Dietz-Band, J., Leversha, M., Affara, N.A., Moore, A.T., Tolmie, J., Schinzel, A., Lerman, M.I., Ferguson-Smith, M.A., Maher, E.R., 1994. Molecular genetic analysis of the 3p − syndrome. Hum. Mol. Genet. 3, 903–908.

Pierpont, M.E.M., Markwald, R.R., Lin, A.E., 2000. Genetic aspects of atrioventricular septal defects. Am. J. Med. Genet. 97, 289–296.

Pras, E., Bakhan, T., Levy-Nissenbaum, E., Lahat, H., Assia, E.I., Garzozi, H.J., Kastner, D.L., Goldman, B., Frydman, M., 2001. A gene causing autosomal recessive cataract maps to the short arm of chromosome 3. Isr. Med. Assoc. J. 3, 559–562.

Ramer, J.C., Ladda, R.L., Frankel, C., 1989. Two infants with del(3)(p25p-ter) and a review of previously reported cases. Am. J. Med. Genet. 33, 108–112.

Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G.G., Stuart, D., 1995. The structure of a Ca^{+2}-binding epidermal growth factor-like domain: Its role in protein-protein interactions. Cell 82, 131–141.

Redegeld, F.A., Caldwell, C.C., Sitkovsky, M.V., 1999. Ecto-protein kinases: Ecto-domain phosphorylation as a novel target for pharmacological manipulation? Trends Pharmacol. Sci. 20, 453–459.

Sekine, M., Nagata, H., Tsuji, S., Hirai, Y., Fujimoto, S., Hatae, M.,

Kobayashi, I., Fujii, T., Nagata, I., Ushijima, K., Obata, K., Suzuki, M., Yoshinaga, M., Umesaki, N., Satoh, S., Enomoto, T., Motoyama, S., Tanaka, K., 2001. Localization of a novel susceptibility gene for familial ovarian cancer to chromosome 3p22–p25. Hum. Mol. Genet. 10, 1421–1429.

Simsir, A., Palacios, D., Linehan, W.M., Merino, M.J., Abati, A., 2001. Detection of loss of heterozygosity at chromosome 3p25–26 in primary and metastatic ovarian clear-cell carcinoma: Utilization of microdissection and polymerase chain reaction in archival tissues. Diagn. Cytopathol. 24, 328–332.

Timmers, C., Whitney, M.A., Thayer, M., Reifsteck, C., Olson, S., Smith, L., Jakobs, P., Leach, R., Naylor, S., Drabkin, H.A., Grompe, M., 1996. Refined mapping of the Fanconi anemia group D complementing gene to a 10cm interval in chromosome 3p25.3. Am. J. Hum. Genet. 59, A238.

Tschentscher, F., Prescher, G., Horsman, D.E., White, V.A., Rieder, H., Anastassiou, G., Schilling, H., Bornfeld, N., Bartz-Schmidt, K.U., Horsthemke, B., Lohmann, D.R., Zeschnigk, M., 2001. Partial deletions of the long and short arm of chromosome 3 point to two tumor suppressor genes in uveal melanoma. Cancer Res. 61, 3439–3442.

von Heijne, G., Gavel, Y., 1988. Topogenic signals in integral membrane proteins. Eur. J. Biochem. 174, 671–678.

Wistuba, I.I., Behrens, C., Virmani, A.K., Mele, G., Milchgrub, S., Girard, L., Fondon 3rd, J.W., Garner, H.R., McKay, B., Latif, F., Lerman, M.I., Lam, S., Gazdar, A.F., Minna, J.D., 2000. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res. 60, 1949–1960.