AMERICAN UNIVERSITY OF BEIRUT

A NOVEL ROLE FOR *CSRP1* IN A LEBANESE FAMILY WITH BOTH CARDIAC DEFECTS AND POLYDACTYLY

by
AMINA ALI KAMAR

A dissertation
submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
to the Department of Biology
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
July 31, 2017

AMERICAN UNIVERSITY OF BEIRUT

A NOVEL ROLE FOR *CSRP1* IN A LEBANESE FAMILY WITH BOTH CARDIAC DEFECTS AND POLYDACTYLY

by
AMINA ALI KAMAR

Approved by:

Dr. Elias Baydoun, Professor — Advisor
Biology

Dr. Georges Nemer, Professor — Co-Advisor
Biochemistry and Molecular Genetics

Dr. Mona Nemer, Professor — Chair of the Committee
Biochemistry, Microbiology and Immunology, University of Ottawa

Dr. Noël Ghanem, Assistant Professor — Member of Committee
Biology

Dr. Zakaria Kambris, Assistant Professor — Member of Committee
Biology

Date of thesis/dissertation defense: July 31, 2017

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: _____Kamar_____Amina_____Ali_____
Last First Middle

◯ Master's Thesis ◯ Master's Project ◉ Doctoral Dissertation

☑ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes
after:
One ---- year from the date of submission of my thesis, dissertation, or project.
Two ---- years from the date of submission of my thesis, dissertation, or project.
Three ---- years from the date of submission of my thesis, dissertation, or project.

Amina 16/8/2017

Signature Date

ACKNOWLEDGMENTS

First, I would like to express my deepest gratitude to my advisor Professor Elias Baydoun, who provided me the opportunity to join the Biology Department as a PhD student. I would like to thank him for his instructions, guidance, scientific advices and encouragement. Without his precious support it would not be possible to conduct this research and complete this degree.

I would like to express my heartfelt gratitude to my co-advisor Professor Georges Nemer who has been my mentor since the master's degree. I am very thankful for his continuous support during my doctoral studies and its related research, for his patience, motivation, and massive knowledge. His guidance helped me in all the time of research and writing of this thesis. I want to thank him also for supporting me personally. He has been extremely caring and patient, especially during my pregnancies and the delivery of my kids.

I have been extremely lucky. I could not have imagined having better PhD mentors. I have learned a lot and benefited from your keen scientific insights. I will be forever indebted to both of you and promise to do my best to make you proud.

Besides my advisors, I would like to thank the rest of my thesis committee: I want to acknowledge Professor Mona Nemer who came all the way through from Canada to attend my defense. I want to thank Dr. Noel Ghanem and Dr. Zakaria Kambris for their insightful comments and thesis revision, but also for the questions which incented me to widen my research from various perspectives. I place on record, my sincere thank you to Professor Julnar Usta, for the continuous encouragement.

I thank my good friend, my colleague, and my only lab mate Athar Khalil for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last four years. I want to thank her for her help and support whenever I needed. Working together made the long hours in the lab seem less stressful.

I also take this opportunity to thank all graduate students, instructors, professors and all other members at the Biology Department who helped and supported me. I specially thank my colleagues: Layla Kamareddine, Dana Bazzoun, Ola El Zein, and Manal Fardon for their unforgettable help at the beginning of my PhD. Special thanks for the member of Dr. Ghanem's lab, Carine Jaafar for her appreciated efforts in providing us with the mouse embryos. My sincere thanks go for Malak Al Dayeh who gave time to help through difficult situations. I will not also forget to thank the support I got from my colleagues at DTS: Salam Idriss, Farah Reda, Farah Ghamloush, and all the rest of the members.

My deep thanks are to all those who have encouraged and supported me all the way through.

I am also very thankful to my parents who taught me the value of education, and to my brothers for their encouragement and support. I am also very thankful to my aunt Aida who selflessly stepped in to the care for my children when I was busy. I would also like to acknowledge and thank my in-laws for their continuous encouragement and support especially my sister-in-Law Nada who always cared for me.

Last but not least, I would like to offer my deep and sincere gratitude for my husband Salam for his patience and everlasting encouragement and attention, and for having confidence in me and pushing me to aim high. I am also thankful to my two precious kids (Yara and Jad) for their patience and for the strength they gave me throughout this period of time. I apologize for all the time I had to spend away from you. You were my inspiration! Without you I would not have completed this journey!

AN ABSTRACT OF THE DISSERTATION OF

Amina Ali Kamar for Doctor of Philosophy
Major: Cell and Molecular Biology

Title: A Novel Role for *CSRP1* in a Lebanese Family with Both Cardiac Defects and Polydactyly

Abnormal cardiac development leads to human congenital heart disease (CHD), which is responsible for the vast majority of neonates' death around the world. Although the major cause is still unknown, CHD has several etiologies ranging from genetic variations, environmental factors in addition to other factors. CHD affect various parts of the heart and it is classified into different broad categories including defects in the septa of atria and ventricles.

In this thesis, we describe a Lebanese family composed from the consanguineous marriage between two first-degree cousins. Out of seven conceived children, two died *in utero* at the ages of 6 and 9 months of unknown causes. Of the remaining five children, three have congenital heart disease (ventricular septal defect, atrial septal defect, and patent ductus arteriosus), and four have polydactyly. Targeted exome sequencing of 119 genes, identified a heterozygote duplication of a 14 nucleotide fragment in exon 5 of CSRP1, causing a frameshift mutation at position 154 of the protein. Genotyping all family members with Sanger sequencing showed that this mutation is segregating with the CHD phenotype but not with polydactyly except for the father who has no cardiac problems and yet is positive for the mutation. The variant was not found in 200 exomes of Lebanese origin neither in the gnomAD. CSRP1 encodes a LIM domain protein, shown to be implicated in smooth muscle function. Our *in silico* analysis revealed that p.E154Vfs*99 mutation totally disrupts the second LIM domain. Thus, we hypothesized that this mutation is deleterious and is the cause of the cardiac defects in the family. First of all, we showed by immunohistochemistry a strong expression of the protein in the heart of mouse embryos at all stages of development starting as early as E12.5 and onwards with a strong nuclear expression in all cardiac compartments, but not in the valves. After generating the MUT CSRP1, immune-staining done on Hela cells shows that p.E154Vfs*99 mutation does not affect the cellular localization of the protein and both the WT and MUT CSRP1 shuttle between the cytoplasm and nucleus in these cells. In addition, by luciferase assay we show that the WT CSRP1 slightly activates the transcription of cardiac developmental promoters NPPA, VEGF, and NOS3. However, the p.E154fs*99 CSRP1 variant significantly abrogates this activation. In addition, we show by co-immunoprecipitation assay that it differentially inhibits the physical association of CSRP1 with SRF and GATA4. We also show that TBX5, a cardiac transcription factor, is a novel CSRP1 partner. Since we have not found any variations in CSRP1 partners, and since the father was normal but still carries the CSRP1 variant, it was very important to perform whole exome sequencing (WES). Although our efforts were unsuccessful in finding a potential

variant linked to polydactyly, we revealed a novel missense mutation in (p.R311S) in *TRPS1*. This variation is inherited from the healthy mother, but still segregates with CHD. Thus, we hypothesized that *TRPS1* variant has no effect unless it is expressed with *CSRP1* variant. *TRPS1* gene encodes a zinc finger transcription factor in which mutations and deletions in TRPS1 cause tricho-rhino-phalangeal syndrome (TRPS) that is characterized by abnormal hair growth and skeletal deformities. In addition, patients with TRPS exhibit a wide range of CHD. Through co-immunoprecipitation, we show that CSRP1 and TRPS1 can physically interact. However, the mutations in each disrupt this physical interaction. Thus, we suggest that the novel frameshift mutation in CSRP1 along with another novel missense mutation in TRPS1 cause the underlying cardiac defects in our family and we propose a digenic model of inheritance to better understand CHD.

The new findings and results presented in this PhD dissertation provide fundamental knowledge into the molecular basis of congenital heart defects and could help unravel additional pathways involved in cardiac development.

CONTENTS

Page

ACKNOWLEDGEMENTS v

ABSTRACT vi

LIST OF ILLUSTRATIONS xiii

LIST OF ABBREVIATIONS xv

LIST OF TABLES xviii

Chapter

I. INTRODUCTION..... 1

A. Molecular basis of cardiac morphogenesis..... 1

1. Signaling pathways in heart development 2
2. Cardiac transcription factors .....
3. Zinc- finger GATA transcription factors 5
a. Cardiac GATA factors 7
i. Cardiac GATA factors expression..... 8
ii. The role of cardiac GATA factors..... 8
b. The GATA-Type zinc finger TRSP1 10

B. Congenital heart disease (CHD) 13

1. Congenital heart disease categories 16
a. Atrial Septal Defect (ASD)..... 17
b. Ventricular septal defect (VSD) 20
c. Patent Ductus Arteriosus (PDA)..... 21
d. Patent Foramen Ovale (PFO) 22
2. Etiology of CHD 23
3. Genetics of CHD..... 25

C. Congenital heart and limb defects 29

1. Holt Oram Syndrome (HOS) 30

a. T-Box Transcription Factor 5 (TBX5) 32
b. TBX5 expressions and Functions 32
2. Tabatznik's syndrome 37
3. Polydactyly and CHD 39

D. LIM domain Proteins 42

1. LIM Domain sequence and structure 44
a. The Family of Cysteine Rich Proteins (CRP) 45
b. Cysteine Rich Protein 1 (CSRP1/ CRP1) 48
i. Regulation of CSRP1 expression 50
ii. Role and expression of CSRP proteins in the heart 51
iii. The dual expression of CSRP proteins in the cell 54

E. Significance of the study 56

II. AIMS OF THE PROJECT 58

III. MATERIALS AND METHODS 61

A. Patient recruitment and clinical examination 61

B. Genetic analysis 61

C. DNA Sanger-Sequencing 62

D. Cell lines 62

E. Site-directed mutagenesis 62

1. Primer phosphorylation 62
2. Polymerase Chain Reaction (PCR) 63
3. Gel electrophoresis 63
4. PCR product purification 63
5. Ligation 64
6. Transformation in bacteria 64

F. Expression vectors and reporter genes 65

G. Antibodies 65

H. Transfection assays to assess CSRP1 target gene promoters 66

I. Protein overexpression 66

J. Nuclear protein extraction 67

K. Protein quantification 68

L. Western blot 68

M. Coimmunoprecipitation (Co-IP) 69

N. Immunostaining 69

O. Immunohistochemistry 70

P. Statistical analysis 71

Q. Antibodies 72

IV. A NOVEL POTENTIAL ROLE FOR *CSRP1* IN A LEBANESE FAMILY WITH BOTH CARDIAC DEFECTS AND POLYDACTYLY: A POTENTIAL CARDIAC DIGENIC EFFECT IN CONJUNCTION WITH *TRPS1* ... 73

A. Abstract 73

B. Introduction 74

C. Materials and methods 77

1. Patient recruitment and clinical examination 77
2. Genetic analysis 77
3. Cell lines and plasmids 78
4. Antibodies 79
5. Transfection assays to assess CSRP1 target gene promoters 80
6. Protein overexpression and western blotting 80
7. Coimmunoprecipitation 81
8. Immunostaining 82
9. Immunohistochemistry 83
10. Statistical analysis 84

D. Results 84

1. CHD Targeted Sequencing: The CSRP1 variant 85
2. Cardiac expression and cellular localization of CSRP1 88
3. Characterization of the transcriptional activity of the CSRP1 mutated protein 89
4. The p.E154Vfs*99 mutation alters the physical and functional

interaction between CSRP1 and SRF 93
5. CSRP1/GATA4 interaction altered by the p.E154Vfs*99 CSRP1 variant 95
6. TBX5 is a novel partner for CSRP1 95
7. Exome sequencing: A novel TRPS1 variant..... 100
8. Role of the digenic CSRP1/TRPS1 variants..... 101

E. Discussion 104

1. CSRP1: A specific integrator of cardiac-enriched transcription factors..... 105
2. CSRP1/TRPS1: A new digenic paradigm in CHD 107
3. Conclusion 109

V. GENERAL DISCUSSION 110

A. p.E154Vfs*99 CSRP1: a deleterious variant 111

B. CSRP1 expression in the heart: A phenotype-Genotype Correlation 113

C. Incomplete Penetrance: Exome and Genome sequencing should solve it.....115

D. Monozygotic Twins: the failure of the genetic code 117

E. Polydactyly: A distinct phenotype requiring different approaches 118

F. Limitations and drawbacks.....119

Appendix

A. SUPPLEMENTARY FIGURES 123

B. SUPPLEMENTARY TABLES.....125

BIBLIOGRAPHY 126

ILLUSTRATIONS

Figure Page

1. Schematic representations of the key events of human and mouse heart development 2

2. The expression of GATA factors in several organs throughout vertebrate development 6

3. The amino acid sequences and structural domains of GATA4, 5, and 6 transcription factors 7

4. A drawing showing the different structural abnormalities detected in CHD 7

5. A normal heart versus a heart with atrial septal defect (ASD) 19

6. Heart diagram representing the types and locations of VSDs 21

7. Schematic representation of patent foramen ovale versus a normal heart 23

8. A model representing CHD etiology 25

9. The regulation of gene expression (*ANF* and *cx40*) 35

10. An EVC (Ellis-Van Creveld syndrome) patient with bilateral polydactyly 41

11. Human LIM-domain protein groups 43

12. The structure of the LIM domain 45

13. The three highly conserved proteins (CRP1, 2, and 3) of the chicken CRP multigene family 46

14. LIM1 and LIM2 zinc-binding domains 49

15. Cardiac and blood vessel LIM domain proteins 52

16. The dual localization of LIM domain proteins 54

17. Congenital heart and limb deformities in a large consanguineous Lebanese family 85

18. Sequencing results of the patients with CSRP1 variant and a schematic representation of CSRP1 86

19. Frontal sections of mouse embryos at different developmental stages (E12.5 and E14.5) 89

20. Cellular localization of WT and MUT CSRP1 proteins 90

21. The transcriptional activity of WT and/or MUT CSRP1..... 92

22. The p.E154Vfs*99 mutation abolishes the physical and functional interaction between CSRP1 and SRF 93

23. Selective inhibition of the GATA4 physical interaction with CSRP1 by the p.E154Vfs*99 mutation 96

24. Transcriptional activity of WT and/or MUT CSRP1 along with GATA4 97

25. The effect of p.E154Vfs*99 mutation on the physical and functional interaction between CSRP1 and TBX5 98

26. The physical interaction between TRPS1 and CSRP1 102

ABBREVIATIONS

CHD	Congenital heart disease
TGF	Transforming Growth Factor
Mesp1	Mesoderm Posterior Protein1
ASD	Atrial septal defect
VSD	Ventricular Septal defect
PDA	Patent Ductus Arteriosus
PFO	Patent Foramen Ovale
TRPS	Tricho-Rhino Phalangeal Syndrome
a.a.	Amino acids
CSRP1/CRP1	Cysteine Rich protein 1
E	Embryonic day
AO	Aorta
AVC	Atrioventricular Canal
LV	Left Ventricle
RV	Right Ventricle
LA	Left Atrium
RA	Right Atrium
OFT	Outflow Tract
DMEM	Dulbacco's Modified Eagle Medium
FBS	Fetal Bovine Serum
ECL	Enhanced Chemiluminescence
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase

TBX5	T-Box transcription Factor 5
TAD	Transactivation Domain
NLS	Nuclear Localization Signal
NPPA	Natriuretic Peptide A
VEGF	Vascular Endothelial Growth Factor
NOS-3	Nitric Oxide Synthase 3
SRF	Serum Response Factor
PTA	Persistent Truncus Arteriosus
ZN	Zinc finger
FL	Forelimb
HL	HindLimb
P	Pharyngeal Arch
TA	Tricuspid Atresia
TOF	Tetralogy of Fallot
AVSD	Atrioventricular Septal Defect
CFC1	Cripto, FRL-1, Cryptic Family1
HOS	Holt Oram Syndrome
CNV	Copy Number Variant
ECA	Extra-cardiac abnormalities
UMS	Ulnar Mammary Syndrome
MZ	Monozygotic
Cx40	Connexin 40
ANF	Atrial Natriuretic Factor
OS	Okihiro Syndrome
EVC	Ellis-van Creveld syndrome
HH	Hedgehog

MKKS	Mckusick-Kaufman Syndrome
LHX	LIM Homeobox protein
LMO	LIM-ONLY protein
C	Cysteine
H	Histidine
MLP	Muscle LIM protein
CaM	Calmodulin
CaMK	Calcium/ Calmodulin dependent Protein Kinase
CArG	Cytosine Adenine rich Guanine element (CC(A/T)6GG)
PKC	Protein Kinase C
MRF4	Myogenic Regulator Factor 4
WT	Wild Type
MUT	Mutant
Min	minute
Hr	hour
RT	Room temperature
Sec	second
gnomAD	Genome Aggregation Database
AF	Allele Frequency
BAV	Bicuspid aortic Valve
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

TABLES

Table Page

1. Phenotypes of $GATA^{-/-}$ mice 6
2. The frequency of CHD in Lebanon 15
3. Some results of studies on CHD and consanguinity 15
4. Common syndromes that result from microdeletions and aneuploidy 27
5. Some of the genes that cause various forms of CHDs with their chromosomal locus in humans 29
6. The expression of CRP proteins, their binding partners and their functions 50
7. List of plasmids used 71
8. List of primers used 72
9. Genotype of Family Members for CSRP1 and TRPS1 variants 87

CHAPTER I

INTRODUCTION

A. Molecular basis of cardiac morphogenesis

The heart formation has shown an evolutionary conserved program from invertebrates to vertebrates. The heart is the first functional organ to develop in which it starts to beat and pump blood by week three in humans during embryogenesis. Mammalian heart development begins with the formation of cardiac crescent from the specification of cardiac progenitor cells within the anterior lateral plate mesoderm that condenses to form two crescent shape heart primordia at around day 15 of human embryonic development and embryonic day E7.5 in mouse development. Those include both myocardial and endocardial lineages. By three weeks of human development and E8.5 of mouse development, the bilaterally paired and condensed cardiac precursors move medially to form the primitive heart tube(David J. McCulley & Black, 2012) .The subsequent patterning of this linear heart tube into ventricles and atria along the antero-posterio axis of the embryo is mediated with a differential gene expression profile which is responsible for the heart morphogenesis into four-chambered pump in adults (Horry & Publishers, 2006). These steps include cardiac looping, trabeculation, septation and valve formation as well as the development of the conduction system (figure 1). The embryonic developmental processes are regulated by transcription factors, and for a transcription factor to be directly involved in cardiac formation, it must control morphogenetic processes and should be expressed in the developing cardiac tissues (Horry & Publishers, 2006).

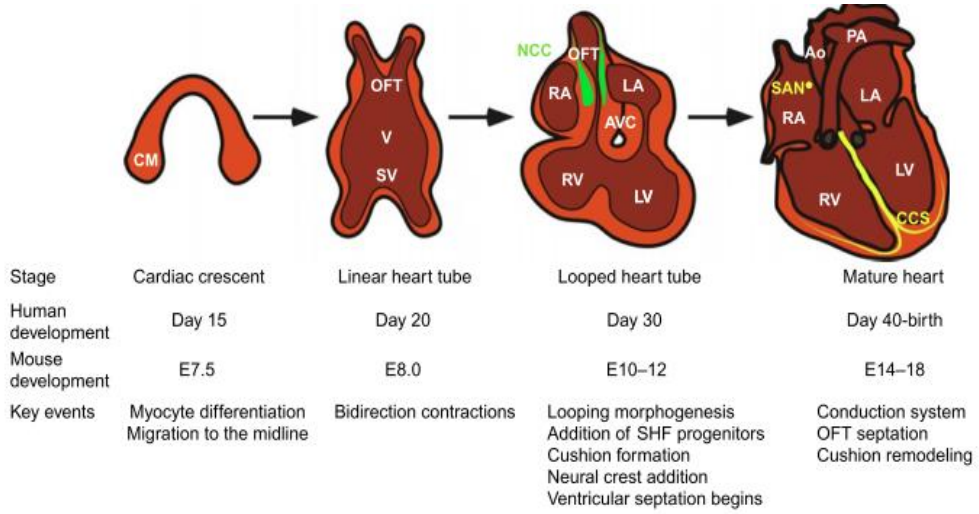


Figure 1. Schematic representations of the key events of human and mouse heart development. Embryonic days of mouse development are represented by (E). Ao, aorta; AVC, atrioventricular canal; CCS, cardiac conduction system; CM, cardiac mesoderm; LA, left atrium; LV, left ventricle; NCC, neural crest cells; OFT, outflow tract; PA, pulmonary artery; RA, right atrium; RV, right ventricle; SAN, sinoatrial node; SV, sinus venosus; V, ventricle(David J. McCulley & Black, 2012).

The differentiation of cardiomyocytes is a complicated multistep phenomenon that needs the proper spatiotemporal integration of many signaling pathways. Previous genetics and embryological reports have detected various signaling pathways that are vital for the initiation of early mesoderm formation and its specification to the cardiomyocyte lineage(Liu & Foley, 2011).

1. Signaling pathways in cardiac development

Cardiomyocyte differentiation is triggered by signals from neighboring tissues. Graded levels of Nodal, a TGFβ- family member, is essential for the specification of mesoderm during the early stages of mesoderm development. However, greater levels of Nodal support the cardiac mesoderm(Brennan et al., 2001). After the cardiac mesoderm specification, Wnt and bone morphogenic protein (BMP) signaling cues are modified during the initial phases of cardiac differentiation. Initially, Wnt signaling

stimulates cardiogenesis; however, it has an inhibitory effect at later stages as the progenitors start to differentiate into several cardiac lineages(Kwon et al., 2007; Naito et al., 2006) . The cardiac differentiation is also promoted by the down regulation of *Isl1* which is required for the induced expansion of cardiac progenitors by Wnt/β-catenin(Kwon et al., 2009).

Although the inhibition of BMP signaling appears significant for the cardiac mesoderm to emerge, yet the function of BMP is not very well defined in the very early phases of cardiac commitment(Yuasa et al., 2005). However, later BMPs appear to be critical for cardiac progenitor's expansions, since the genetic deletion of the BMP receptor in Mesp1-expressing mesoderm causes major defects in heart development (Klaus et al., 2007). In addition, Notch signaling is essential for cardiac precursors since the deletion of Notch1 in the second heart field (SHF) increases the number of cardiac precursors along with the increase in Wnt signaling activity which promotes the expansion of cardiac progenitors(Kwon et al., 2009).

Retinoic acid (RA) which is a derivative of vitamin A, in addition to its derivatives (retinoids), are involved in regulating several processes during cardiogenesis through binding to nuclear RA receptors (RAR) and RA X receptors (RXR). Deregulated retinoid signaling lead to defects in cardiac formation(Zaffran, Robrini, & Bertrand, 2014). Earlier studies established that RA deficiency causes a reduction in posterior cardiac *Tbx5*$^{+}$ progenitors that were observed at the 10-somite stage and later in mouse(Niederreither et al., 2001). Sirbu *et al.* demonstrated that loss of the posterior *Tbx5* expression befalls by the 3- somite stage. In addition, gene knock out studies in mouse have revealed central roles for *Isl1* (Cai et al., 2003) and *Tbx5*(Benoit G. Bruneau et al., 2001) in heart tube development along the anteroposterior axis. Trunk

mesoderm tissue-specific *retinaldehyde dehydrogenase* is encoded by *Raldh2*. *Raldh2* knockout mice embryos suffered from aberrations in cardiac morphogenesis due to elevations in *Isl1*$^{+}$ progenitors subsequent to increased *FGF8* signaling. It was suggested that RA limits the number of cells attaining an *Isl1*$^{+}$ fate by restricting *Fgf8* expression, therefore putting RA close the topmost of a significant cardiac signaling hierarchy: RA -| FGF8 → Erk1/2 → Isl1. Thus, it was concluded that *Raldh2* represses *Fgf8* in the posterior cardiac area and that a mutual antagonism may occur between *RA* and *FGF8* signaling throughout several stages of cardiac morphogenesis. Thus, RA-FGF8 antagonism disruption during heart tube formation may be the cause of molecular defects in some forms of congenital heart disease(Sirbu, Zhao, & Duester, 2008).

2. Cardiac transcription factors

The mature heart is formed during late phases of embryogenesis after a series of complicated morphogenetic events that depend on the functional presence of several proteins (G. Nemer & Nemer, 2001). The regulation of each step in cardiogenesis requires a proper spatio-temporal expression of cardiac transcription factors. Chamber specific expression of some genes and the gradient expression of others are regulated by combinatorial interaction between these transcription factors. Thus, normal cardiac function requires a fine-tuned cooperative interaction between diverse classes of transcription factors and aberrations in function and expression lead to cardiac malformations (G Nemer & Nemer, 2001; Srivastava, 2001). However, the molecular mechanisms that are necessary for the spatio-temporal regulation of cardiac transcription continue to be unclear. The evolutionary conservation of protein-protein interactions from drosophila to mammals and the fact that point mutations that interrupt

these interactions generate developmental cardiac defects in humans and mice, reveal the significance of protein-protein interactions for proper cardiac development (M. Nemer, 2008).

The homeodomain protein Nkx2-5, GATA family zinc finger proteins including GATA4,5, and 6, SRF (MADS box proteins) and MEF2 factors, T-box factors including Tbx1,2,5,18, and 20 in addition to the Lim homeodomain protein Is11 are a group of transcription factors that are essential for cardiac development (B. L. Black & Cripps, 2010; Garg et al., 2003b; Greulich, Rudat, & Kispert, 2011; Harvey, 2002; A. He, Kong, Ma, & Pu, 2011; Peterkin, Gibson, Loose, & Patient, 2005). These basic transcription factors interact either with each other or with other set of transcription factors to regulate heart formation.

3. Zinc- finger GATA transcription factors

GATA proteins belong to a large and evolutionary conserved family of zinc domain transcription factors that are characterized by the presence of a zinc finger as a major component of their structure. GATA transcription factors can bind the DNA sequence A/G GATA A/T since they harbor a highly conserved DNA-binding domain(Molkentin, 2000). They are divided into two subfamilies based on their structure and sequence homology in which they execute key non-redundant functions in vertebrates and invertebrates (Malik et al., 2001). The first subfamily includes GATA1,2, and 3 which are highly expressed in hematopoietic cells and are required for the differentiation of central nervous system, while the second subfamily includes GATA4,5, and 6. The latter are expressed in several mesoderm and endoderm derived-tissues including the heart, lungs, liver and gut(Lentjes et al., 2016) (figure 2).

Exceptionally, GATA2 acts as an essential regulator of transcription in endothelial cells(Linnemann et al., 2011). The zinc finger motifs are more than 70% identical amongst the six GATA factors, however, the N- and C-terminal domain sequences harbor less similarity between them(Morrisey, Ip, Tang, & Parmacek, 1997). GATA factors are considered to be essential for development since single *Gata* knockout mice face embryonic lethality(Lentjes et al., 2016) (Table.1). For example, the knockout of GATA1, 2, and 3 in mice lead to embryonic lethality because of hematological defects(Simon, 1995).

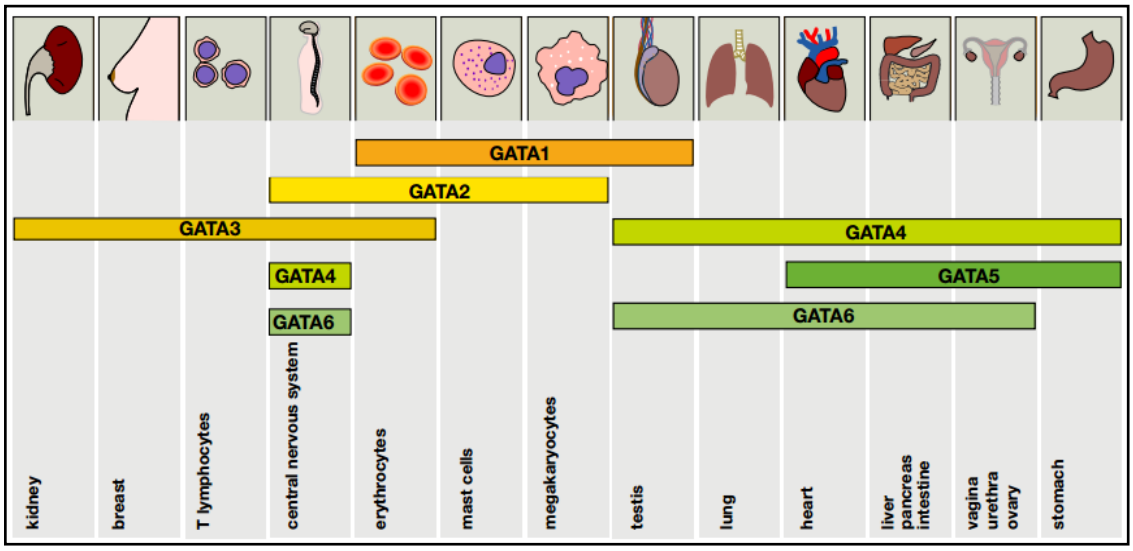


Figure 2. The expression of GATA factors in several organs throughout vertebrate development. The expression of all GATA factors is displayed in the corresponding tissues in which the distribution of the expression patterns roughly reveals the two GATA subfamilies (GATA1/2/3 against GATA4/5/6) (adapted from Lentjes et al., 2016).

Table 1. Phenotypes of GATA$^{-/-}$ mice.(adapted from Lentjes et al., 2016)

Name	Phenotype (embryonic day)	Abnormality	Reference
GATA1	die (11.5–12.5 dpc)	Defective erythroid cell maturation	Fujiwara *et al.*
GATA2	die (12.5 dpc)	Severe anaemia	Tsai *et al.*
GATA3	die (11–12 dpc)	Massive internal bleeding and severe deformities of the brain and spinal cord	Pandolfi *et al.*
GATA4	die (9.5 dpc)	Defects of heart morphogenesis and ventral closure of the forgut	Kuo *et al.*
GATA5	Viable and fertile	Females exhibited pronounced genitourinary abnormalities that included vaginal and uterine defects and hypospadias	Molkentin *et al.*
GATA6	die (5.5–7.5 dpc)	defects of visceral endoderm function and subsequent extra-embryonic endoderm	Morrisey *et al.*

a. Cardiac GATA Factors

GATA4, 5 and 6 expression takes places in the mesodermal precursor that develop into the heart. The molecular weights of mouse GATA4, 5, and 6 are 48, 45, and 42 kDa respectively. Their amino acid sequences are ~ 85% identical within the DNA-binding domain containing zinc finger and basic domains (Jeffery D. Molkentin, 2000). Within the DNA binding region, mouse *GATA-4* is ~70% identical to mouse *GATA-1* and *Drosophila pannier* , this suggests that the sequence of GATA family members is highly conserved throughout evolution (figure 3)(Jeffery D. Molkentin, 2000).

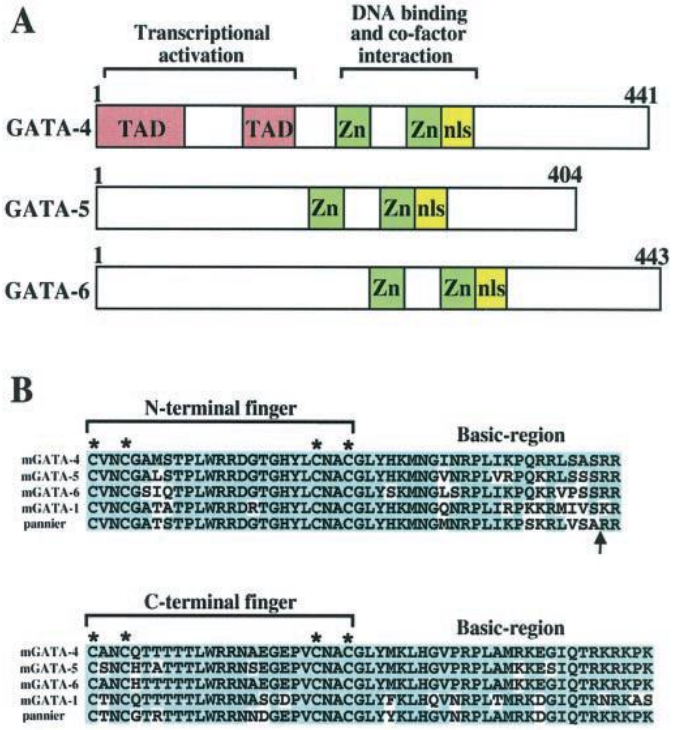


Figure 3. The amino acid sequences and structural domains of GATA-4,-5, and -6 transcription factors. A. The DNA binding domain of GATA-4 contains two distinct zinc finger motifs (Zn) and a nuclear localization sequence (nls) at the C-terminus in which they constitute the DNA-binding and protein-protein interaction domains. At the N-terminus, GATA-4 also comprises two transcriptional activation domains (TAD). **B.** amino acid sequence of the two zinc finger domains and the adjacent basic regions. For comparison, Mouse GATA-1 and *D. pannier* amino acid sequences are also presented. The unshaded areas reveal difference in amino acid sequence while the blue shaded region displays identity to GATA-4 (consensus). Asterisk represents the cysteines that constitute each zinc finger subdomain, while the arrow designates the location of a 3-amino acid insertion that is exclusively present in *pannier*; sequence not shown (Molkentin, 2000).

i. The expression of Cardiac GATA factors

Cardiac GATA factors including GATA4,5, and 6 reveal differential expression and regulation during development (M. Nemer, 2008) with GATA4 being the predominant transcript in cardiomyocytes at all stages (Georges Nemer & Nemer, 2003). Specifically, GATA4 and 6 are expressed in the myocardial layer of the developing heart and are the strongest at the posterior end, while GATA5 myocardial expression is restricted to few atrial cells and completely disappear at E14.5(Georges Nemer & Nemer, 2010). Thus, GATA5 expression is totally restricted to endocardial cells (M. Nemer, 2008).

ii. The Functions of Cardiac GATA Factors

So far, GATA4 is the most extensively studied member of the family. It was initially isolated as an upstream regulator of *NPPB*, the gene that encodes the brain natriuretic peptide BNP (Grépin et al., 1994). GATA4 has appeared as the nuclear effector of several signaling pathways which modulate its function through post-translational modifications and protein-protein interactions. Several studies have shown that GATA4 plays an essential role in cardiac development and is critical for survival of the embryo(Peterkin et al., 2005; Pikkarainen, Tokola, Kerkelä, & Ruskoaho, 2004). For example, *Gata4* knock-out mouse embryos show *cardiac bifida* and hypoplastic ventricles (Jeffery D. Molkentin, Lin, Duncan, & Olson, 1997), and the deletion of GATA4 from cardiac cells caused myocardial thinning which reveals its vital role in cardiomyocytes proliferation(Pu, Ishiwata, Juraszek, Ma, & Izumo, 2004; Zeisberg et al., 2005). In addition, GATA4 acts in combination with core cardiac transcription factors to regulate heart development such as its interaction with Nkx2-5 (Liberatore, Searcy-Schrick, Vincent, & Yutzey, 2002; Lien et al., 1999; Searcy, Vincent,

Liberatore, & Yutzey, 1998). GATA4 and NKx2-5 were reported as disease causing genes in atrial septal defects(Hirayama-Yamada et al., 2005). Moreover, the expression of GATA4, Nkx2-5 and SRF were found to be essential for the activation of embryonic cardiac α-actin gene(Sepulveda, Vlahopoulos, Iyer, Belaguli, & Schwartz, 2002).

Several studies have shown that GATA4, 5 and 6 bind to similar consensus sequences in cardiac (Charron, Paradis, Bronchain, Nemer, & Nemer, 1999; J D Molkentin, Kalvakolanu, & Markham, 1994; G Nemer, Qureshi, Malo, & Nemer, 1999) and some gastric promoters (Gao et al. , 1998; Tamura et al., 1993), and thus enhancing their activity. Therefore, a possible functional redundancy among these proteins is raised, in particular with the overlapping of GATA4 and GATA6 expressions in cardiomyocytes. It was widely speculated that the two proteins may be able to compensate for each other in heart development (M. Nemer, 2008). Yet, the exact role of GATA6 is not fully elucidated although many lines of evidence suggest that it participates in cardiovascular development (G. Nemer & Nemer, 2010). During murine embryonic and postnatal development, GATA6 is highly expressed in vascular smooth muscle cells. At embryonic stage, *Gata6* knockout mice die due to abnormalities in the extraembryonic endoderm (Morrisey et al., 1998). In addition, tissue-specific GATA6 deletion in neural crest-derived smooth muscle cells lead to a persistent truncus arteriosus (PTA) and an interrupted aortic arch, a phenotype associated with a severely attenuated expression of a signaling molecule essential for vascular and neuronal patterning, semaphoring 3C. These findings suggest that GATA6 is needed for appropriate patterning of the aortic arch arteries (Lepore et al., 2006). *Wnt2* has been identified by microarray assay analysis as a GATA6 target gene in smooth muscle cells and cardiomyocytes after GATA6 transient overexpression. In addition, GATA6 has

been interestingly detected as Wnt2 target in which they both form a transcriptional loop that regulates posterior heart development(Tian et al., 2011). GATA6 mutations have been related to CHD etiology. For instance, two GATA6 mutations were detected in PTA patients in which the transcriptional activity of GATA6 on downstream target genes involved in the development of cardiac outflow tract was disrupted (Kodo et al., 2009). Zhao *et al.* have showed that knock out of GATA6 from embryos generated from embryonic stem cells by tetraploid embryo complementation show no detectable change in cardiac morphogenesis or in the expression of cardiac genes and develop till E10.5. However, mice lacking one copy of GATA 4 and 6 are lethal at the embryonic stage suggesting that both transcription factors can compensate for each other during cardiac development (Xin et al., 2006).

GATA5 has a specialized role in endocardial development due to its high expression in endocardial cells. GATA5 mutants in *zebrafish* have *cardiac bifida* and lack endocardial cells, reduced expression of Nkx2-5 cardiac markers in addition to hypoplastic ventricles (Reiter et al., 1999). However, the in activation of GATA5 in mice did not produce any cardiac phenotype since GATA4 compensates for the lack of GATA5 in these cells(Jeffery D. Molkentin, 2000) . As a conclusion, the GATA4,5, and 6 transcription factors share similar related functions during cardiovascular development in which defects cause CHD and other heart disorders(Lentjes et al., 2016).

b. <u>The GATA-Type zinc finger TRSP1</u>

GATA factors are characterized by the presence of two zinc fingers Cys-X2-Cys-X17- Cys-X2-Cys, ZNI and ZNII (Lentjes et al., 2016; Jeffery D. Molkentin,

2000), while invertebrates i.e. *Drosophila melanogaster* and *Caenorhabditis elegans* express a class of important proteins that harbor one GATA-zinc finger. The C-terminal zinc finger (ZNII) is present in both vertebrates and non-vertebrates demonstrating that ZNI was duplicated from ZNII (C. He, Cheng, & Zhou, 2007).

TRSP1, an atypical member of GATA proteins(Kunath, Lüdecke, & Vortkamo, 2002), is a vertebrate protein that is highly conserved between *Xenopus* and mammals and yet contains a single GATA-type zinc finger(Malik et al., 2001; Momeni et al., 2000). Kaiser at al. identified TRPS1 gene on human chromosome 8q24.1 which encodes 1281 a.a. protein with a calculated molecular weight of 141.6 kDa.

TRSP1 contains nine putative zinc finger domains (eight of which are different) with the seventh finger representing the GATA-type while zinc fingers 8 and 9 reveal homology to a conserved domain of lymphoid transcription factors that belong to Ikaros family(Malik et al., 2001; Momeni et al., 2000).The Ikaros family conserved domain is known to mediate repression of transcription through protein- protein interactions(Malik et al., 2001). So far, only the functions of motifs 7,8 and 9 have been determined(Kaiser et al., 2004).

Mutations in human *Trsp1* gene is involved in the autosomal dominantly inherited TRP (tricho-rhino-phalangeal) syndromes which are human diseases with skeletal and craniofacial malformations(Kunath et al., 2002; Momeni et al., 2000). Most cases in TRPS are mutations or deletions in TRPS1(Maas et al., 2015a; Malik et al., 2001). Specifically, some of the major features of the TRPSs include hip malformations, sparse scalp hair, bulbous tip of the nose, protruding ears, short stature, brachydactyly, and cone-shaped epiphyses in the phalanges. This demonstrates its essential role in development(Malik et al., 2001; Momeni et al., 2000). Previously, many studies have

focused on the role of TRPS1 in bone and joint development. During embryonic mouse development, *Trsp1* is highly expressed at the cartilage region, developing joints, hair follicles and the developing nose. Moreover, *Trsp1* showed high expression at skeletal condensation regions, lung and gut mesenchyme, intervertebral disks, and trachea in addition to a complex expression pattern in the brain (Kunath et al., 2002).

TRSP1 differs from other GATA proteins by its *in vitro* and *in vivo* activity as a sequence-specific transcriptional repressor rather than an activator since although it binds a GATA sequence, it fails to activate GATA transactivation reporter(Malik et al., 2001). Recent studies have shown that a broad proportion of patients with TRPS display wide range of congenital cardiac defects, varying from minor to severe abnormalities and including persistent foramen ovale, persistent ductus arteriosus, prolapse of cardiac valves, aortic stenosis, anomalous pulmonary venous return, and left cardiac insufficiency (Maas et al., 2015b; Verheij et al., 2009a). In addition, many studies have revealed the role of TRPS1 in joint and bone development(Nomir et al., 2016). Through *in situ hybridization*, Nomir *et al*. revealed a restricted Trps1 expression, which was observed in endocardial cushions of the outflow tract, and in leaflets of all mature cardiac valves. These results suggested that the Trps1 proximal promoter sequence comprises some of the tissue specific Trps1 regulatory region. Moreover, these findings partially explain why patients with TRPS show a broad range of congenital cardiac defects, although Trps1 expression is observed in a more restricted fashion(Nomir et al., 2016). Still, how mutations in TRPS1 lead to abnormal cardiac development is unclear (Nomir et al., 2016).

Many of the same cardiac transcription factors are re-employed to regulate the maturation of heart chamber, development of conduction system, and remodeling of endocardial cushion(Oka, Xu, & Molkentin, 2007; Olson, 2006). Each of the core transcription factors is involved in the regulation of the expression of the others in which they function in a jointly reinforcing transcriptional network (B. Black, 2007; A. He et al., 2011; Olson, 2006). Some of the major factors that are involved in cardiac development also act as biochemical partners. This reflects a complicated molecular and genetic relationship that controls multiple phases of cardiac and conduction system development. Thus, mutations in numerous genes that encode core cardiac transcription factors are related to congenital heart disease (CHD) (David J. McCulley & Black, 2012).

B. Congenital heart disease (CHD)

In humans, heart defects are common congenital malformations (Lin et al., 2012). Congenital heart disease (CHD) is the major cause of mortality and morbidity in the world(D. J McCulley & Black, B, 2012) in which it is recognized in 10%of still births(Fahed, Gelb, Seidman, & Seidman, 2013). Congenital heart defects arise before birth and usually refer to abnormalities in the heart's structure or function (Benoit G Bruneau, 2008). CHD accounts for one third of all main congenital defects and it is considered as the most frequent type of birth anomalies. The number of infants born with CHD worldwide each year is 1.35 million. In addition, CHD has a variable prevalence crosswise countries and continents. It occurs in 8.1 per 1000 live births in North America. However, the prevalence of CHD in Asia is 9.3 live births. This disparity in prevalence is due to expanded proportions of parental consanguinity(Fahed

et al., 2013). Parental consanguinity leads to a two to three fold elevation in the risk for a wide range of CHD phenotypes, as reported in Saudi Arabia(Becker, Al Halees, Molina, & Paterson, 2001), Lebanon (Nabulsi et al., 2003), and South India(Ramegowda & Ramachandra, 2006). Many studies were done and supported an association between consanguineous parentage and congenital heart disease(Shieh, Bittles, & Hudgins, 2012). In the Arab world, the population is characterized by having high consanguinity rate which may lead to a higher incidence and prevalence of CHD than in the rest of the world. This makes CHD of greater concern in the Arab region (Aburawi et al., 2015). Lebanon is one of the Middle-East countries that suffer of high levels of consanguinity. Bitar *et al.* showed that in Lebanon, the frequency of infants born with CHD between 1980 and 1995 was 11.5 per 1000 live births as compared to other countries of the world (Bitar et al., 1999) (Table 2). This ratio is higher than what was reported in literature due to higher rates of consanguinity (Nabulsi et al., 2003). Additionally, the percentage of CHD is also greater in the children of first cousin couples than in off springs born to couples with far consanguinity (Aburawi et al., 2015; Chehab et al., 2007). First cousin marriages (where the individuals share 1/8 of their genes) are very common in some cultures and could lead to higher disease risk(Shieh et al., 2012). However, the level of consanguinity varies with religion, urban, rural characteristics, and with socio-economic conditions(Chehab et al., 2007). Nabulsi et al. showed that 20% of 759 Lebanese patients with various forms of congenital heart defects were found to be from first degree cousin mating (Nabulsi et al., 2003)(Table 3).

Table 2. The frequency of CHD in Lebanon. This table shows the frequency of CHD disease in Lebanon between 1980 and 1990 (red box) as compared to different parts of the world based on literature(adapted from Ramegowda & Ramachandra, 2005).

Country/City	Year	Frequency/ 1000	References
England	1957-1971	6-9	Bound et al,
Liverpool	1960-1969	6.6	Kenna et al,
Prague	1977-1984	6.64	Samanek et al,
Lebanon	1980-1995	11.5	Bitar et al,
Bohemia	1980-1990	6.61	Samanek et al,
Western Australia	1980-1989	7.65	Bower and Ramsay
NSW and ACT2-Australia	1981-1984	4.3	Kidd et al
Qatar	1984-1994	12.23	Robida et al,
Norway	1987-1990	10.2	Meberg et al,
Karachi	1987-1992	4	Hassan et al,
Central Australia	1993-2000	17.5	Bolisetty et al,
South Australia	1993-2000	12	Annual report 2000
Oman	1994-1996	7.1	Subramanyan et al,
Bosnia	1994-1999	6.12	Begic et al,
Egypt	1995-1996	1.01	Bassili et al,
Shimla	1995	2.25	Thakur et al,
Saudi Arabia	1997-2000	2.4	Alabdulgader
New Delhi	2001	4.2	Chadha et al,

Table 3. Some results of studies on CHD and consanguinity (Shieh et al., 2012).

		No. subjects		Percent with consanguinity			
Study	**Country**	**CHD**	**Controls**	**CHD**	**Controls**	**Reported statistics**	
Becker et al. 2001	Saudi Arabia	891	3212	40.4%[a]	28.4%	Z statistic	P<0.001
Nabulsi et al. 2003	Lebanon	759	19,589	20.2%[a]	13.2%	X^2	P<0.0001
Roodpeyma et al. 2002	Iran	346	346	22%	19.1%	X^2	NS
Ramegowda et al. 2006	India	144	200	40.3%	15.5%	[b]	P=0.0001
Yunis et al. 2006	Lebanon	173	865	17.9%[a]	9%	X^2	P<0.001
Chehab et al. 2007	Lebanon	1585	1979	19.4%[a]	14.4%	X^2	P<0.0001
Bassili et al. 2000	Egypt	894	894	44.1%	23.8%	[c]	

[a]First-cousin consanguinity
[b]Data not available
[c]Average inbreeding coefficient 0.021 in CHD cases versus 0.011 in controls (P<0.05)

1. Congenital heart disease categories

Congenital heart disease affects most parts of the heart and can be classified into three broad categories: cyanotic heart disease, left-sided obstruction defects, and septation defects. Infants with cyanotic heart disease appear blue as a result of the mixing of oxygenated and deoxygenated blood. The third main type of congenital heart disease which is septation (accounts for 30% of CHDs) can affect the atrial septum formation (atrial septation defects, ASDs), ventricular septum formation (ventricular septal defects, VSDs) or development of structures in the middle part of the heart (atrioventricular septal defects, AVSDs). Bicuspid aortic valve (BAV) and patent ductus arteriosus (PDA) are other types of congenital malformations that do not neatly fit into the three major categories. The most common congenital heart disease is BAV, while septation defects are considered the next common(McDaniel, 2001). The most frequently occurring CHDs encounter septal and vulvular defects (Hoffman, Kaplan, & Liberthson, 2004). Their severity ranges from moderately slight subclinical defects such as patent foramen oval (PFO) to more complicated defects such as Tetralogy of Fallot (TOF) or triscupid atresia (TA) that can threaten one's life(M. Nemer, 2008) (figure 4).

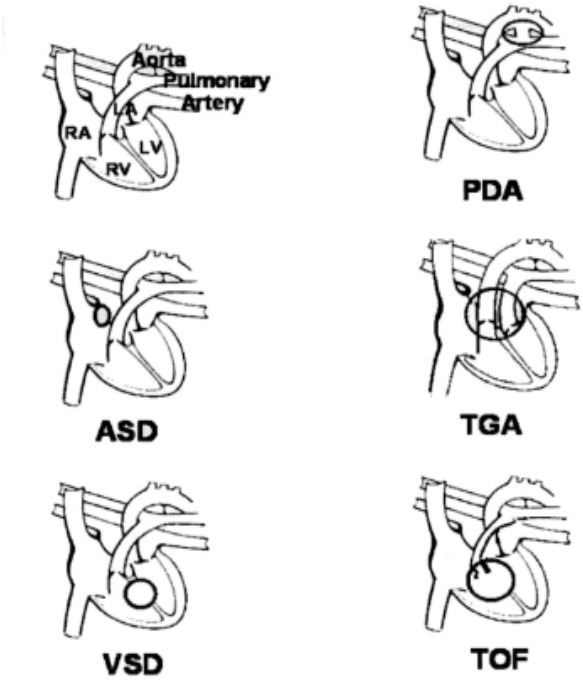


Figure 4. A drawing showing the different structural abnormalities detected in CHD. Those include: Transposition of Great Artery (TGA), Atrial Septal Defect (ASD), Ventricular Septal Defect (VSD), Patent Ductus Arteriosis (PDA), and Tetralogy of Fallot (TOF). Right Atrium (RA), Right ventricle (RV), Left ventricle (LV)(Adapted from Nemer G., 2006a) .

The early diagnosis of CHDs is essential for prevention of later more complicated conditions such as cardiovascular as heart failure, valve replacement, or stroke. By far, most CHDs have a familial history, yet, sporadic and isolated cases of CHDs are not uncommon (M. Nemer, 2008).

a. Atrial Septal Defect (ASD)

A septal defect, which arises due to abnormal development of the atrial septum in humans, is a hole in different parts of the atrial septum which allows some amount of blood to enter from the left to right atrium instead of flowing into the left ventricle (figure 5). This causes a kind of tenacious communication between the left and right atria resulting in right heart failure during the mid-years of life. ASDs account for about

10 % of CHD. ASD occurs two to three times more in females than in males, and it can arise in healthy children as well (McDaniel, 2001). The hole could be either single or multiple and might be located anywhere in the atrial septum, and thus, according to this ASDs are categorized into 4 key types: 1) Ostium secundum (or Fossa ovalis, 75% of all ASDs), 2) Ostium primum and 3) Sinus venosus (10% of ASD) and 4) Unroofed coronary sinus (Harvey, 2002; McDaniel, 2001; Pathways, 2016; Ramegowda & Ramachandra, 2005). Fossa ovalis is located in the central part of the septum. Sinus venosus are present in the upper part of the septum specifically at the joining part with the superior vena cava. Although ostium primum is classified as an ASD, it is more likely considered an atrioventricular septal defect (AVSD). It is less frequent than fossa ovalis and is usually associated with Down syndrome(Geva, Martins, & Wald, 2014). Sinus venosus is a defect that affects the communication between one or more of the right pulmonary veins and the cardiac end of the superior vena cava (superior vena cava type) or the posterior-inferior atrial wall just above the inferior vena cava-right atrial junction. The most common location of the defect (around 87%) is between the right upper pulmonary vein and the superior vena cava, resulting from deficiency of the tissue that separates these two veins. Unroofed coronary sinus is considered as uncommon atrial communication which results from partial or complete unroofing of the tissue separating the coronary sinus from the left atrium. This permits a shunt through the defect and the coronary sinus orifice. Raghib syndrome is a result of an association of a coronary sinus septal defect and persistent left superior vena cava(Geva et al., 2014).

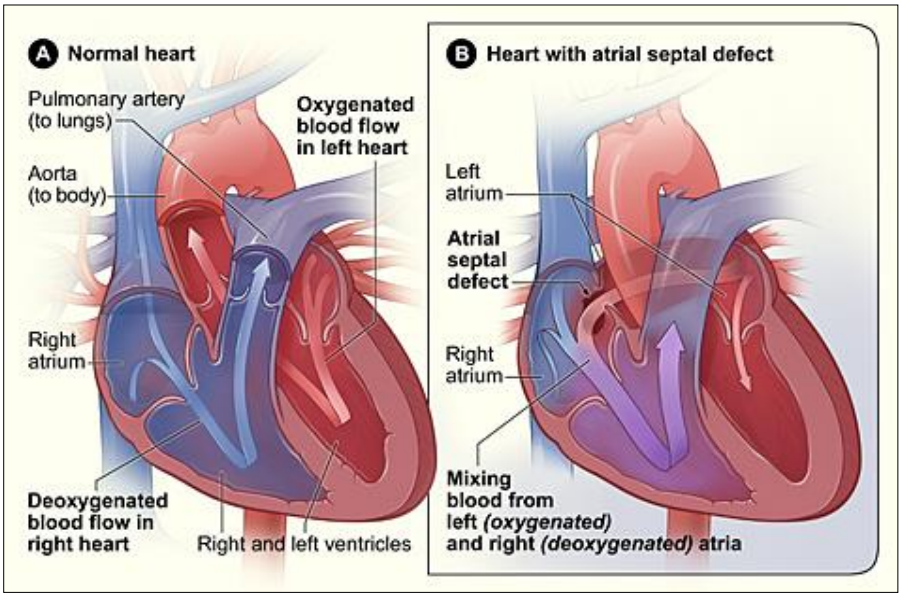


Figure 5. A normal heart versus a heart with atrial septal defect (ASD). (A) The structure and blood flow inside a normal heart. **(B)** A heart with an atrial septal defect. The hole allows oxygen-rich blood from the left atrium to mix with oxygen-poor blood from the right atrium (https://www.nhlbi.nih.gov/health/health-topics/topics/holes/types).

Some of the ASDs forms have some identified genetic causes. HOS patients have limb defects in association with ASDs (Lyons et al., 1995). This syndrome is due to mutations in TBX5, a T-box transcription factor. In addition, mutations in the gene encoding NKx2.5, which is a TBX5 partner, have been associated to sporadic and familial ASD cases(Abu-Issa, Smyth, Smoak, Yamamura, & Meyers, 2002; Goldmuntz, Geiger, & Benson, 2001). Mutations in GATA4 were also found in familial cases with ASDs which reveals out the main role of GATA4 in septation in which both NKx2-5 and TBX5 are GATA4 partners. This cooperative activity defines a network that contributes to ASDs. Other genes including *Cited2* and *Fog2* transcription factor genes have been suggested to be as disease-causing genes including ASDs in humans though mutations have not been identified in humans(Garg et al., 2003b; L. Li et al., 1997; G Nemer & Nemer, 2001; Svensson et al., 2000). The understanding of ASDs will be

more elucidated through finding downstream targets of these particular transcription proteins(Nemer G., 2006b) .

b. Ventricular septal defect

VSD is a defect or hole in the septum (wall) between the heart's ventricles ("Ventricular septal defect.," 2009). It is the most common defect(Nemer G., 2006b) in which they account for 20 % of congenital heart defects (McDaniel, 2001). The ratio of males and females having VSD is equal (Pathways, 2016). The most common CHD is VSD involving the interventricular septum(Bitar, Diab, Sabbagh, Siblini, & Obeid, 2001; Hoffman & Kaplan, 2002). Although VSD occurs as an isolated condition, yet, it can occur along with other conditions of CHD such as TOF, TA, TGA and many other defects(Nemer G., 2006b) . VSDs are divided into four types: 1) Perimembranous, 2) supracristal (subpulmonary or subaortic), 3) inlet (VSD of the atrioventricular canal), and 4) muscular(Pathways, 2016) (figure 6). Those forms of VSDs result from the failure of the formation of the ventricular septum. A perimembranous VSD is found in the area of the membranous septum which places it in the outflow tract of the left ventricle and directly below the aortic valve. A supracristal VSD is positioned superior to the crista supraventricularis which thus put it in the right ventricular outflow tract and immediately below the right cusp of the aortic valve. An inlet VSD, which is a type of endocardial cushion defect, is present posteriorly in the septum just inferior to the tricuspid and mitral valves. Finally, a muscular VSD is located in the muscular ventricular septum(Pathways, 2016).

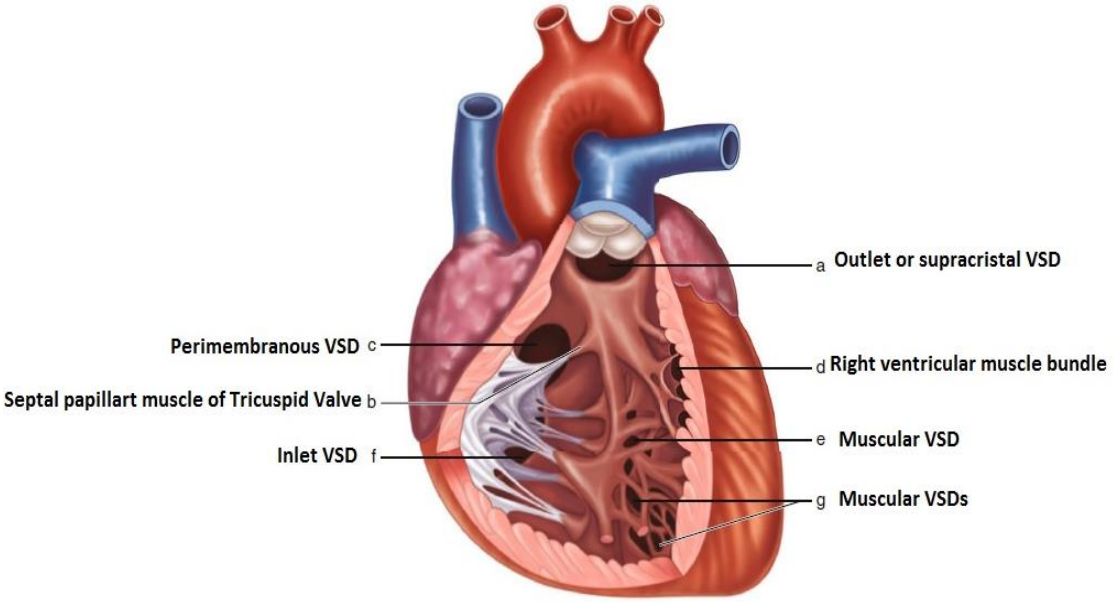


Figure 6. Heart diagram representing the types and locations of VSDs. Since the anterior surface of the heart has been removed, the ventricular septum is revealed as observed from the right side Adapted from (adapted from Pathways, 2016).

The etiology of VSD is heterogeneous and so complicated. Chromosomal aberrations including structural variations and aneuploidy in addition to sporadic mutations in different genes are associated with VSD. DiGeorge and Holt-Oram syndromes are two examples of well-defined syndromes with known genetic cause in which patients have VSD. Mutations in genes encoding cardiac transcription factors including GATA4, NKx2-5 in addition to signaling molecules such as CFC1 have been often present in VSD cases(Pathways, 2016).

c. Patent Ductus Arteriosus (PDA)

In the fetus, the ductus arteriosus appears normally as a wide clear vessel that connects the descending aorta and the pulmonary trunk (Nemer G., 2006b). PDA condition arises when the ductus arteries fail to close within 72 hours of birth. PDA is

one of the most common congenital heart defects in which it causes up to 30 % mortality and morbidity rates in infants (Dice & Bhatia, 2007). It occurs more in females than in males. 12q24 locus was found to be potentially linked to PDA(Yin et al., 2002). This locus includes various cardiac genes such as Tbx5 and Shp2 that are implicated in HOS and Noonan syndromes, respectively (Mani et al., 2002; Tartaglia et al., 2002). The inactivation of many genes in mice, that encode transcription factors, lead to PDA (Nemer G., 2006b).

d. Patent Foramen Ovale (PFO)

Patent foramen ovale (PFO) is an opening between the primum and secundum atrial septa located at the fossa ovalis which persists after one year of age (figure 7). Normally during fetal life, the foramen ovale is essential for the communication between the left and right atria and helps circulate blood more quickly in the absence of lungs. Thus, blood can migrate from the veins to the right side of the fetus and travel to the left side of the heart without having to pass through the lungs. Although its prevalence decreases with age, PFO is present in 20-30% of the population and occurs in 1 out of 4 individuals. In most cases, it shows no threat to one's life. However, under some hemodynamic conditions, PFOs can open and function as a duct that allows thrombi to pass from the systemic venous circulation to the systemic arterial circulation. This can theoretically cause a stroke. In majority of cases, PFO have no symptoms (Calvert, Rana, Kydd, & Shapiro, 2011). Still, cyanosis could happen if the baby has both PFO and other heart condition("Patent foramen ovale," 2015).
Improvements in imaging technology have induced the enhancement of diagnostic accuracy (Donofrio et al., 2014).

Numerous assemblages of various defects can originate together, and can lead to more intricate CHD in which Tetralogy of Fallot (TOF) is the most frequent form (Silversides, Kiess, et al., 2010; Silversides, Salehian, et al., 2010).Thus, understanding the developmental mechanisms that underlie the formation of cardiac septal defect has significant implications for the diagnosis, prevention and treatment of congenital heart disease.

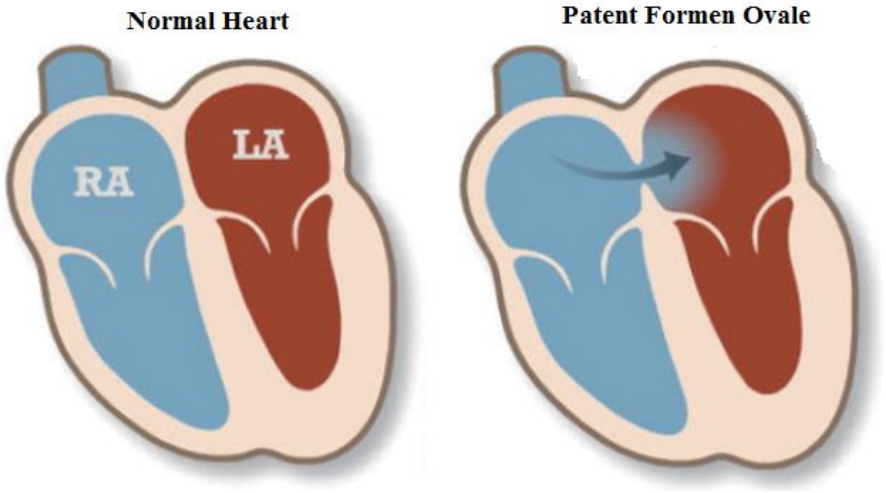


Figure 7. Schematic representation of patent foramen Ovale versus a normal heart. The arrow shows of the PFO. LA = left atrium; RA = right atrium(adapted from Bang et al., 2015).

2. Etiology of CHD

In general, most of congenital heart anomalies are sporadic(Ramegowda & Ramachandra, 2005). Although the major cause is still unknown, CHD has several etiologies that are often divided into genetic and non-genetic groups which range from genetic variations, environmental factors in addition to other factors (Lage et al., 2012) (figure 8). Around 80% of CHD is considered multifactorial and arises through several consolidations of genetic and environmental factors. Approximately, 20% of the cases

can be associated to chromosomal abnormalities, Mendelian syndromes, non-syndromal single gene complaints or teratogens. Patients with CHD commonly have Down syndrome and velocardiofacial syndrome. Non-syndromal forms of CHD have been linked to at least 30 genes. Their contribution to CHD remains unknown but is presumed to be relatively small (Blue, Kirk, Sholler, Harvey, & Winlaw, 2012). There is limited evidence for the contribution of specific environmental factors to the cause of CHD. However, there are some factors that might reduce the risk of CHD in infants such as folic acid supplementation before and after conception period, completing rubella vaccination before pregnancy, and sustaining good glycemic control in diabetic mothers(Blue et al., 2012).

The evidence of the genetic involvement to the etiology of CHD is based on a cognate re-occurrence risk 2 to 4 times greater than the total incidence(Loffredo et al., 2004), the common recurrence of CHD in off springs with partial or complete aneuploidy(Van Karnebeek & Hennekam, 1999), the presence of various micro-deletion syndromes that encompass CHD including DiGeorges [MIM 188400], Williams [MIM 194050] and Alagille [MIM 118540] syndromes, and the authentication of point mutations in numerous genes that are commonly involved in cardiac development (Goldmuntz et al., 2001; C. Li et al., 2013; Y. Q. Yang, Li, Wang, Liu, Chen, Zhang, Wang, Jiang, et al., 2012).

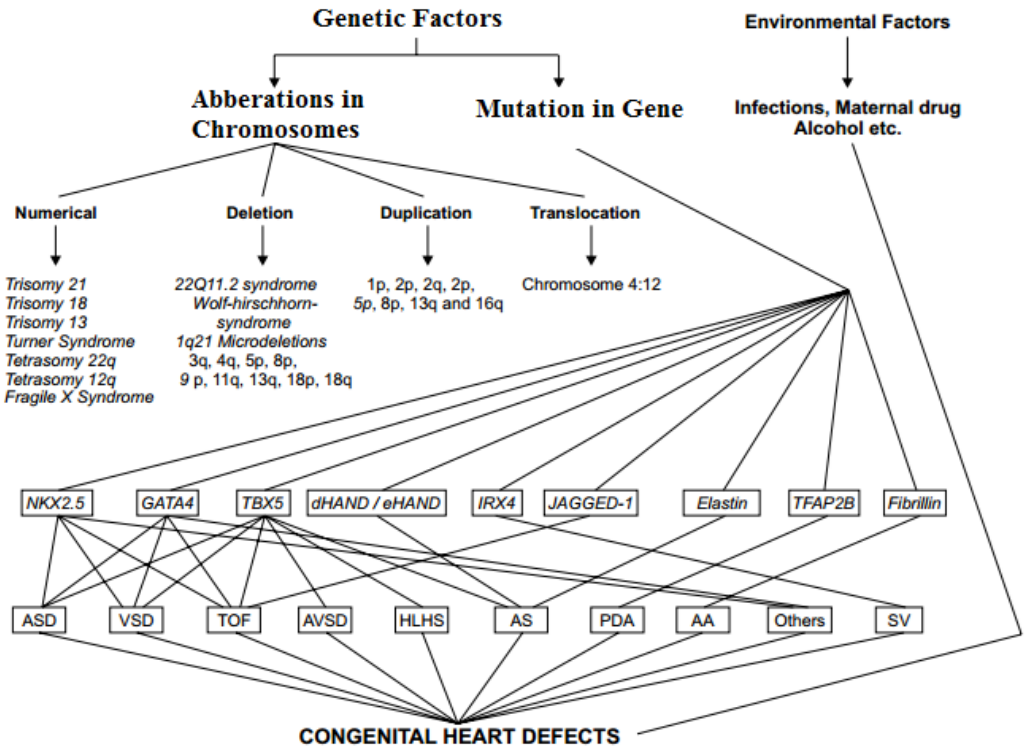


Figure 8. A model for CHD etiology. CHD is caused by either genetic factors (including chromosomal aberrations or gene mutation) or environmental factors. The model also explains how a single gene accounts for various forms of CHD phenotypes, how a single CHD is caused by the association of diverse genes, and how the interaction of variable genes produces CHDs (Adapted from Ramegowda & Ramachandra, 2005).

3. Genetics of CHD

CHD is a genetically heterogeneous disease that is associated with mutations in various developmental genes and chromosomal aberrations(Postma, Bezzina, & Christoffels, 2015). Familial mutations in CHD can be either autosomal recessive, autosomal dominant, or X-linked. Those mutations are expressed with high penetrance and with different manifestations (Fahed et al., 2013). Molecules involved in heart development are highly interconnected since mutations in different genes cause identical malformations. In addition, the range from which heart abnormalities that

emerge for an identical gene mutation reveals genomic context, maternal-fetal environment, cardiac biomechanics, and other factors are essential consequences that influence the clinical effects of CHD mutations (Fahed et al., 2013). Generally, individuals suffering of CHD are at higher risk for extra-cardiac congenital abnormalities and for neurodevelopmental complications (Egbe, Uppu, Lee, Ho, & Srivastava, 2014; B. S. Marino et al., 2012).A range of 25 to 40 % of CHD cases occur in association with other anomalies or a known syndrome. However, the major number of infants born with CHD do not have other birth anomalies(Richards & Garg, 2010).

About 30% of infants with inborn CHD have chromosomal abnormalities (Pierpont et al., 2007). A significant proportion of CHD is accounted for aneuploidy which is defined as abnormal chromosomal number. About 50 % of cases (i.e. 1/600 births)(Fahed et al., 2013) that are born with trisomy 21, have CHD that ranges from atrial and ventricular septal defects to atrioventricular canal abrasions. This percentage increases to 80% in Trisomy 13 in which heterotaxy and laterality deformities become more frequent. In addition, in Trisomy 18 approximately all individuals will bear CHD most frequently in the form of septal defects. In Turner syndrome or monosomy X, about one third of females have CHD. Usually, the defects are commonly on the left side of the heart structures. The most well recognized diagnoses are biscupid aortic valve, aortic stenosis, hypoplastic left heart syndrome, and coartation of the aorta(Richards & Garg, 2010). The incidence of CHD in males with Klinefelter syndrome, or 47, XXY is 50%, in which atrial septal defects and patent ductus arteriosus are the most common(Pierpont et al., 2007). On the other hand, a few forms of CHD are linked with long-realized chromosomal aberrations such as trisomy of

chromosome 21 and chromosome 22q11 deletion. These genotype–phenotype observations reveal an essential conclusion which is that cardiac malformations are due to altered dose of specific genes and not because of a global variation in genomic content (Fahed et al., 2013) (Table 4).

Table 4. Common syndromes that result from microdeletions and aneuploidy. (Richards & Garg, 2010)

Syndrome	Cardiac Anomalies	% with CHD	Other Clinical Features
Trisomy 13	ASD, VSD, PDA, HLHS	80%	Microcephaly, holoprosencephaly, scalp defects, severe mental retardation, polydactyly, cleft lip or palate, genitourinary abnormalities, omphalocele, microphthalmia
Trisomy 18	ASD, VSD, PDA, TOF, DORV, CoA, BAV	90-100%	Polyhydramnios, rocker-bottom feet, hypertonia, biliary atresia, severe mental retardation, diaphragmatic hernia, omphalocele
Trisomy 21 (Down syndrome)	ASD, VSD, AVSD, TOF	40-50%	Hypotonia, developmental delay, palmar crease, epicanthal folds
Monosomy X (Turner Syndrome)	CoA, BAV, AS, HLHS	25-35%	Short stature, shield chest with widely spaced nipples, webbed neck, lymphedema, primary amenorrhea
47, XXY (Klinefelter Syndrome)	PDA, ASD, mitral valve prolapse	50%	Tall stature, hypoplastic testes, delayed puberty, variable developmental delay
22q11.2 deletion (DiGeorge Syndrome)	IAA Type B, aortic arch anomalies, truncus arteriosus, TOF	75%	Thymic and parathyroid hypoplasia, immunodeficiency, low-set ears, hypocalcemia, speech and learning disorders, renal anomalies
7q11.23 deletion (Williams-Beuren Syndrome)	Supravalvar AS, PPS	50-85%	Infantile hypercalcemia, elfin facies, social personality, developmental delay, joint contractures, hearing loss

ASD, atrial septal defect; VSD, ventricular septal defect; PDA, patent ductus arteriosus; HLHS, hypoplastic left heart syndrome; TOF, tetralogy of Fallot; DORV, double outlet right ventricle; CoA, coarctation of aorta; BAV, bicuspid aortic valve; AVSD, atrioventricular septal defect; IAA, interrupted aortic arch; AS, aortic stenosis; PPS, peripheral pulmonic stenosis.

Copy number variation (CNV) is defined as submicroscopic deletions or duplications of chromosomes which has emerged as essential CHD contributors (Costain, Silversides, & Bassett, 2016). Although karyotyping is genome-wide, its resolution is low and it rarely recognizes structural defects which are less than 5-10 Mb in size. Karyotyping reveals chromosomal abnormalities that include duplication or

deletion of hundreds of genes in addition to the expression of complicated developmental phenotypes(Costain et al., 2016). CNV has identified captious genes which are dosage sensitive and critical for cardiac morphogenesis. 22q11.2 deletion syndrome and Williams–Beuren syndromes are genetic disorders that were first identified with common CNVs linked to highly penetrant CHDs. A role for CNVs in CHDs with extracardiac abnormalities (ECAs) and in isolated CHDs has been recently supported through research inquiries and clinical diagnostic testing. Approximately 2-3% of CHDs with ECAs and 3-10% of isolated CHDs are associated with CNVs. It is suggested that there is an elevation in large, infrequent genic CNVs in CHDs patients. This indicates that the load of CNV can play an essential role in disease diagnosis and gene discovery(Lander & Ware, 2014). Rare *de novo* and inherited copy number variants (CNVs) have been shown to appear in 5-10% probands with CHD(Breckpot et al., 2010, 2011; Erdogan et al., 2008; Hitz et al., 2012; Lage et al., 2012; Lalani et al., 2012; Silversides et al., 2012; Soemedi et al., 2012). This evidence marks the great heterogeneity between genetic factors involved in CHD. The 1-3 Mb 22q11.2 deletion that leads to *DiGeorges* velocardio-facial syndrome has been the only single common cause recognized(Warburton et al., 2014).

So far, at least 50 human disease genes have been associated with CHD(Anderson, Spicer, Brown, & Mohun, 2014), however, small set of developmental genes (for example, NKX2.5 (Shott et al., 1998), GATA4 (Garg et al., 2003b) and NOTCH1(Garg et al., 2005) and TBX5(Q. Y. Li et al., 1997) form the majority of CHD-associated mutations (Table 5).

Table 5. Some of the genes that cause various forms of CHDs with their chromosomal locus in humans(Ramegowda & Ramachandra, 2005)**.**

Gene	Chromosome locus	Type of defects
CSX/ NKX2.5	5q34	ASD, VSD, AV block, TOF, Ebstein malformations and Tricuspid valve abnormalities
GATA4	8p22-23	ASD,VSD, AVSD, Pulmonary valve thickenings
TBX5	12q24.1	ASD, VSD, AVSD,TOF, HLHS, AS
dHAND/eHAND	4q33 and 5q33 respectively	AS
IRX4	5p15.3	SV
JAGGED-1	20p12	TOF
Elastin	7q11	AS
TFAP2B	6p12	PDA
Fibrillin	15q21	AA

While the underlying genetic basis for many of these defects remains elusive, mutations in genes encoding core cardiac transcription factors have emerged as major contributors to many forms of CHD (David J. McCulley & Black, 2012).

C. Congenital heart and limb defects

The patterning of heart and the limbs is a multistep procedure that includes respectively the heart and limb fields' specification, the formation of well-defined signaling regions that notify cells of their positions, elucidation of positional signals, and growth and differentiation regulation(Krause et al., 2004). Congenital cardiac and limb defects often happen in association with each other and are categorized as heart-hand syndromes. It remains unclear whether the heredity of heart-hand syndromes protrudes from common or different genetic disorders despite the numerous similarities

they share in clinical manifestations. The diagnosis and control of complicated group of autosomal dominant heart-hand syndromes with largely variable phenotypic expression will be improved through the elucidation of the genetic etiologies. The frequency of the association of congenital heart disease with radial defects is high and has been identified since the case reported by Birch- Jensen in 1949(Silengo, Biagioli, Guala, Lopez-Bell, & Lala, 1990).

1. Holt Oram Syndrome (HOS)

Holt Oram Syndrome (MIM No. 142900) is the classical heart-hand disorder characterized by cardiac septal defects in addition to radial ray limb abnormalities (McKusick, 1992). HOS is a rare autosomal dominant human syndrome that is caused by *TBX5* mutations(C T Basson et al., 1999; Q. Y. Li et al., 1997). In 1960, Holt and Oram who first described HOS observed atrial septal defects associated with congenital anomaly of the thumb in members of 4 generations of a family. Hundreds of studies have described the clinical feature of HOS which revealed to be variable, but with complete penetrance. Up to 95% of the patients with HOS have cardiac malformation and all of them have upper limb anomaly. As a conclusion, atrial septal defects and ventricular septal defects are the most prevalent cardiac defects in HOS. Other defects could range from first degree heart blockage (conduction defects) to multiple structural defects(B G Bruneau et al., 1999) .

The strategy for a precise diagnosis of HOS is the uniform existence of upper-limb radial ray deformities, which could be symmetrical or asymmetrical (even unilateral) irrespective of the incidence or lack of cardiovascular disease (Pierpont et al., 2007). Other limb malformations including syndactyly of digits other than the thumb, polydactyly, or lower-limb defects, craniofacial abnormalities, and/or sign of noncardiac

visceral organ anomalies (such as heterotaxy) make HOS unlikely(Allanson & Newbury-Ecob, 2003; Craig T Basson et al., 1995; Mcdermott et al., 2005; Newbury-Ecob, Leanage, Raebum, & Young, 1996).

Upper limb anomalies affect the upper limbs exclusively (hand, wrist, and/or arm defects)(C J Hatcher & Basson, 2001); no lower limb abnormalities have been reported. These abnormalities are often symmetrical and are always bilateral, mainly involving the radial ray. The thumb which is the most affected structure can be hypoplastic, triphalangeal, or completely absent. Such abnormalities could be either minor or severe such as phocomelia in which patients have malformed or missing limbs(Huang, 2002). Other congenital malformations were reported along with cardiac defects and upper limb abnormalities. For example, few cases with congenital cardiac defects and conduction disorder has been reported in patients with ulnar-mammary syndrome (UMS), a disease characterized by apocrine abnormalities and posterior (ulnar) malformations of upper limbs(Linden et al., 2009; Meneghini et al., 2006). In 2014, Bogarapu *et al.* reported a case with features of both HOS and UMS(Bogarapu et al., 2014). The patient has bilateral symmetric limb malformation, congenital heart defects and cardiac conduction disease. This phenotype is due to a contiguous deletion of TBX5 and TBX3, thus expanding the spectrum of the disease(Bogarapu et al., 2014).

There is an overlap in clinical features between HOS, Townes-Brocks syndrome (caused by *SALL1* mutations) and Okihiro syndrome (OS). Okihiro syndrome which has been seen in several reports refers to familial incidence of radial sided hand defects in association with Duane syndrome of eye retraction. OS was shown to be caused by mutations in *SALL4* gene(Kohlhase et al., 2002).

a. T-Box Transcription Factor 5 (TBX5)

Tbx5 is a member of T-box family that encodes transcription factors which have a highly conserved DNA binding motif (T-box)(Craig T Basson et al., 1995; Q. Y. Li et al., 1997). Members of this family are found in both vertebrates and invertebrates(Craig T. Basson et al., 1997). Research studies done on frog, zebrafish and newt revealed that proper levels of Tbx5 gene expression is necessary for cardiogenesis and limb identity (Cathy J. Hatcher & Basson, 2001) . Tbx5 contributes to many aspects of cardiovascular development in which it is expressed in all cardiac compartments specifically higher levels are present in the atrial chambers(Cathy J. Hatcher & Basson, 2001; Liberatore et al., 2002).

b. TBX5 expressions and Functions

Tbx5 is considered to be the initial determinant of vertebrate cardiac growth and has been involved in cardiomyocyte proliferation regulation(C J Hatcher & Basson, 2001). Initially, mouse *Tbx5* is expressed in a contagious region that corresponds to cardiac crescent and forelimb field (B G Bruneau et al., 1999). Also, *zebrafish* cells expressing Tbx5 are originally organized in two bilateral strips which divide into an anterior group of cells which contribute to the cardiac primordium and a posterior cluster of cells which migrate to the future pectoral fin bud(Ahn, Kourakis, Rohde, Silver, & Ho, 2002).The correlation between Heart/Limb defects in Mendelian syndromes in humans suggests that there is a coordination between heart and limbs and the presence of cardiomelic field (Wilson, 1998). In humans, mutations in Tbx5 would lead to Holt-Oram syndrome (HOS). By studying multiple unrelated families, Tbx5, the

disease causing gene, was linked to chromosome 12q2(Craig T. Basson et al., 1994; Bonnet et al., 1994). Mutations in Tbx5 gene were demonstrated in various affected individuals and families(C T Basson et al., 1999; Craig T. Basson et al., 1997). The T-box domain, which is the DNA binding region, is highly conserved across species and among members of the T-box family of transcription factors. TBX5 consists of nine coding exons. At least two alternatively spliced isoforms modify the coding region to add or remove the terminal exon, whose presence modifies TBX5 activity but is not necessarily required(C T Basson et al., 1999; Ghosh et al., 2001).

However, different studies done on families with disparate phenotypes of Holt-Oram syndrome revealed heterogeneous mutations in TBX5. It was found that mutations that create null alleles cause extensive heart and limb abnormalities. However, missense mutations of TBX5 can produce dissimilar phenotypes. For example, one class of missense mutations causes severe cardiac anomalies but only minor limb abnormalities. Other missense mutation can cause the opposite phenotype (Huang, 2002).

The mutation detection rate in TBX5 coding region can range from 20%-50% in familial cases and up to 40% in cases with negative family history. The mutations are distributed along the gene with few hot spots (codon 273) and nucleotide 824 (Ghosh *et al*., 2001). HOS is probably caused by haploinsufficiency of TBX5 and most mutations lead to premature termination of protein product. Genotype-phenotype correlations have been studied. In one of the studies, patients with missense mutations had normal arms and absent or hypoplastic thumb, while those with frameshift mutation had severe phocomelia. However, missense mutations in TBX5 may produce distinct phenotypes. Mutations at aminoacid 237 cause minor cardiac malformation, but severe skeletal

deformities. Significant cardiac defects, but only minor skeletal abnormalities are caused with G80R mutation. What is also interesting is that intrafamilial variations do exist(C T Basson et al., 1999) . Two families with distinct mutations showed an intrafamilial pattern and interfamilial variations as reported by Basson *et al.* Thus, it seems that TBX5 mutations serve as gross tune and that other factors including the genetic background act as a finer tuning for phenotypic expression of HOS. A frameshift mutation in TBX5 was reported in a pair of identical monozygotic (MZ) twins. This mutation produced a truncated protein at aminoacid 263. Both twin had ASD, VSD, radial club and delayed carpal ossification. However, the twins' hands were not identical. This suggests that the genetic background alone cannot explain the conflicting features in those identical twins. In other words, there are other factors that affect the phenotypic expression of HOS(Huang, 2002).

It is clear now that the elucidation of the TBX5 intracellular pathway might allow studying the causes of common cardiac and limb abnormalities. It was demonstrated that several genes are regulated by TBX5 through analyzing *Tbx5* knock-out mice. Those included *ANF* and connexin 40 (Cx40). Mouse models of HOS ($TBX5^{del/+}$ heterozygote) show dramatic reduced expression of *cx40* (connexion 40), a gene essential for the conduction of electrical impulses throughout the heart, and therefore accounting for the conduction defects seen in HOS(Benoit G. Bruneau et al., 2001).The expression of ANF was also reduced in mice expressing 50% Tbx5 levels. The expression of cx40 and ANF was specifically shown to be regulated through the physical interaction between Tbx5 and Nkx2.5 (figure 9). These studies provide one potential explanation for Holt-Oram syndrome conduction system defects, suggest mechanisms for intrafamilial phenotypic variability, and account for related cardiac

malformations caused by other transcription factor mutations(Benoit G. Bruneau et al., 2001). On the other hand, Kraus et al identified a new chicken PDZ-LIM protein called LMP4, which is expressed in all organs that express Tbx5 and-4 including the developing eyes, heart, forelimbs and hind limbs. The latter can interact with the transactivation domain at the C-terminal of Tbx5 and -4 transcription factors(B G Bruneau et al., 1999; Krause et al., 2004).Thus, factors other than TBX5 mutation itself and genetic background might be contributing to this phenotypic variability.

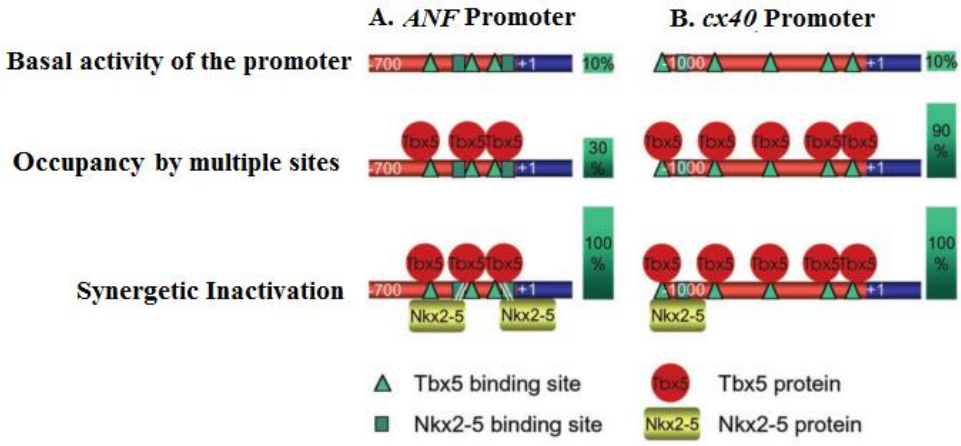


Figure 9. The regulation of gene expression (*ANF* and *cx40*) by two cardiac transcription factors, Tbx5 and Nkx2-5. (A) Regulation of the *ANF* gene. **(B)** Regulation of the *cx40* gene. *ANF* harbors three Tbx5-binding sites (presented as triangles) and two Nkx2.5–binding sites (presented as squares). *Cx40* harbors five Tbx5-binding sites and only one Nkx2.5–binding site. The number of Tbx5-binding sites reveals the percentage of activation of each gene by Tbx5, with *cx40* more dependent on full occupancy of its Tbx5-binding sites. Addition of physical interactions between Nkx2-5 and its DNA target as well as with Tbx5 results in synergistic activation of the target genes, resulting in maximal expression. The models shown are based on previously reported data (Adapted from Benoit G. Bruneau, 2002).

Other specific missense mutations in TBX5 have been shown to disrupt the interaction between TBX5 and GATA4 (a gene known to cause non-syndromic congenital heart disease, particularly septal defects). Therefore, mutations in GATA4

are likely to cause septal defects through its interaction with TBX5(Garg et al., 2003a) . Mutations in SALL4, a zinc finger transcription factor (associated with Duane-radial ray syndrome), cause similar upper limb and heart defects as mutations in TBX5. In addition, limb deformities in Okihiro syndrome (OS), which is due to mutations in SALL4, are identical to those in HOS(Al-Baradie et al., 2002; Jürgen Kohlhase et al., 2002). HOS is usually characterized by the absence of abducens neurons, renal abnormalities, anal stenosis and CHDs most likely VSD, and congenital deafness(J. Kohlhase et al., 2002; Jürgen Kohlhase et al., 2002). About 13% of OS affected individuals with SALL4 mutations have CHDs. OS and HOS are sometimes confused because of their association with CHDs and since SALL4 mutations have been detected in patients who were first diagnosed with HOS(Brassington et al., 2003; J Kohlhase et al., 2003).

In a study done by Koshiba-Takeuchi et al., Tbx5 was shown to regulate Sall4 expression in the forelimb and heart of developing mouse. In addition, mice that were heterozygous for a gene trap allele of Sall4 revealed cardiac and limb defects that resemble human disease. As a conclusion, the positive and negative interaction between Tbx5 and Sall4 is essential to fine tune patterning and morphogenesis of embryonic heart and the anterior forelimb and which provides a common mechanism for heart-hand syndromes(Koshiba-takeuchi et al., 2006). In addition, transcription factor TBX5 was found to interact with NKx2.5 and synergistically activate other genes. Thus, understanding protein-protein interactions in the developing heart can lead to the identification of candidate genes involved in common congenital heart diseases(Huang, 2002).

The morphogenesis and patterning of developing structures involve specific interactions between the regulators of transcription and signaling molecules. Patterning features are usually regulated by transcription factors through either the activation or repression of downstream target genes. The interaction between transcription factors during embryonic heart development can construct the basis for gene dosage sensitivity and is mandatory for the activation of robust genes(Seidman & Seidman, 2002). The key interactions between TBX5, GATA4 and NKX2-5 are thought to be of great importance since when disturbed, cause CHD (Garg et al., 2003b; Hiroi et al., 2001). The transcriptional regulation of the developing embryo complexity is not clearly understood, however, the severe malformations that result from dominant mutations in genes encoding transcription factors in human syndromes reveal its prominence (Seidman & Seidman, 2002). In addition, many features of heart patterning are not elucidated yet regardless of these identified interactions (Koshiba-takeuchi et al., 2006). Thus, searching the human genome databases can help us elucidate candidate genes that might be regulated by TBX5(Huang, 2002).

2. Tabatznik's syndrome

Temtamy and McKusick described a family that was studied by Tabatznik in 1978. In this family, upper limb defect including Brachydactyly type D and congenital cardiac arrythmiasas occurred in three generations as a dominant autosomal or X-linked trait. The latter was called Heart-Hand syndrome II in order to differentiate it from HOS(Silengo et al., 1990). Tabatznik'syndrome Type II (Jakubek et al., 2016; Silengo et al., 1990; Temtamy & McKusick, 1978) is a rare Heart-Hand syndrome which is characterized by upper-limb deformities, and congenital heart arrhythmias has no clear

etiology. Heart-Hand syndrome Type III (McKusick 1992; Ruiz de la Fuente S 1980) (MIM No. 140450), which is characterized by cardiac conduction disease including intraventricular delays and sick sinus syndrome, is phenotypically similar but poorly understood. In Brachydactyly type C, skeletal deformities are restricted to hand and feet. No septal defects have been associated to patients with Heart-Hand syndrome type III(Craig T Basson et al., 1995). There are several clinical resemblances between the heart-hand syndromes and a diversity of less complicated autosomal dominant partial phenocopy conditions such as familial atrial septal defects (ASDs) with conduction disease occurring without limb defects, and familial limb deformities occurring without cardiac malformations (Bizarro & Callahan, 1970; McKusick, 1992). In 1980, La Fuente and Prieto described a new kind of autosomal dominant Heart-Hand syndrome type III. In this case, affected patients had in addition to cardiac arrhythmias, type C brachydactyly. However, there were no upper limb defects in addition to those of hand and feet were detected. The cardiac arrhythmias were composed of sick sinus syndrome and intraventricular conduction abnormalities. In 1990, Lala et al. reported a family in which the 8-year- proband had comparable heart and upper limb defects than the Tabatznik cases. In addition, other features such as facial dysmorphism, cryptorchidism and mild mental retardation were observed. Another study done by Ramos et al. in 1999 reported an 8-year-old girl with a new case of Heart-hand Syndrome Type II. The affected girl had bilateral IV and V petite metatarsals and metacarpals. In addition, she had cardiac arrhythmia composed of abnormal migration of sinusal pacemakers, mild mitral and tricuspid prolapse without clinical symptoms, and some facial dysmorphic features with normal intelligence. The girl's family history revealed the existence of hand defects in some individuals of the maternal side.

3. Polydactyly and CHD

Polydactyly is a condition in which a person has extra fingers and /or toes per hand and/or foot. Polydactyly is the most frequently occurring congenital hand and foot deformation that ranges from minor soft tissue duplications to major bony malformations. Polydactyly has an autosomal dominant inheritance and a prevalence of 5 to 19 per 10000 live births. Hand polydactyly occurs in approximately 1 in 1000 live births while hand/feet polydactyly is rare and has a frequency of 1 in 1000000 individuals(Fish et al., 2016; Hwang et al., 2016). Polydactyly which is almost common, occurs due to abnormal aberrant anterior-posterior patterning(Biesecker, 2012). Since the embryology of limb development is complicated, various phenotypes of polydactyly can be observed(Bev Guo, M.D., Steven K. Lee, M.D., and Nader Paksima, D.O., 2013). Thus, polydactyly can be either pre-axial (radial side), post-axial (ulnar), and central (non-border digit) depending on the location of the extra digit(Bev Guo, M.D., Steven K. Lee, M.D., and Nader Paksima, D.O., 2013). Polydactyly has been reported in other species such as cats, pigs, horses and chicken, it is considered to be ubiquitous(Gorbach, Mote, Totir, Fernando, & Rothschild, 2010). The gene responsible for pre-axial polydactyly as well as complex polysyndactyly is located on chromosome 7q36(Hwang et al., 2016). Polydactyly can also occur with some genetic diseases including CHD. Atrioventricular canal defect (AVCD) is considered a common congenital heart defect (CHD) which represents about 7.5% of all heart malformations. In addition, it is considered secondary to an extracellular matrix anomaly. In about 75% of the cases, AVCD is associated with extracardiac defects.

Polydactyly syndromes including Ellis-van Creveld, short rib polydactyly, oral-facial-digital, Bardet-Biedl and Smith-Lemli-Opitz syndromes are associated with AVCD. Experimental studies done on mice revealed the involvement of the Sonic Hedgehog (SHH) pathway syndromes with postaxial polydactyly and heterotaxia, and ciliary dysfunction as pathomechanism for these defects. However, anatomic variances in AVCD in the diverse groups are perhaps due to different genetic causes(Digilio et al., 2011). For example, Ellis-van Creveld syndrome (EVC) which is a chondral and ectodermal dysplasia characterized by short ribs, polydactyly, growth retardation and ectodermal and heart defects (figure 10). It is a rare disease with approximately 150 cases reported worldwide. The exact prevalence is unknown, but the syndrome seems more common among the Amish community. Prenatal abnormalities (that may be detected by ultrasound examination) include narrow thorax, shortening of long bones, hexadactyly and cardiac defects. After birth, cardinal features are short stature, short ribs, polydactyly, and dysplastic fingernails and teeth. Heart defects, especially abnormalities of atrial septation, occur in about 60% of cases. Cognitive and motor development is normal. This rare condition is inherited as an autosomal recessive trait with variable expression. Mutations of the EVC1 and EVC2 genes, located in a head to head configuration on chromosome 4p16, have been identified as causative (Baujat & Le Merrer, 2007). Patients with EVC syndrome usually have atrioventricular canal defect as a distinctive cardiovascular malformation which is characterized by the prevalence of partial form of atrioventricular canal with common atrium and persistent left superior vena cava(B. Marino & Digilio, 2000). In humans, EVC expression was studied in human embryonic tissue via in situ hybridization. Low levels of EVC expression were detected in developing heart, bone, lung and kidney at Carnegie stages

19 and 21. However, EVC in the bone was shown to be expressed in the developing vertebral bodies, upper and lower limbs in addition to the ribs. Interestingly, EVC expression was also detected in atrial and ventricular myocardium which include both ventricular and interventricular septa(Ruiz-Perez et al., 2000). Molecular studies have shown that EVC1 and EVC2 (LBN) are essential for cilia function. In addition, EVC were found to be essential for HH signaling transduction. Cilia are microtubule-based organelles that protrude the surface of the cell to promote transduction of several developmental signaling pathways. Dysfunction in cilia has been associated in ciliopathies. In addition, the coordinated function between EVC proteins is important for cilia-dependent heart formation. Recently, several studies have suggested CHD as part of ciliopathy disease spectrum and implicated cilia in the etiology of CHD(Shiaulou Yuan, Samir Zaidi, 2014).

Figure 10. An EVC (Ellis-Van Creveld syndrome) patient with bilateral polydactyly. The latter is revealed by short finger(Baujat & Le Merrer, 2007).

Another example on such cases is McKusick–Kaufman syndrome which is caused by mutations in the MKKS gene and has an autosomal recessive mode of inheritance. In addition to polydactyly, patients have heart defects and genital

abnormalities, and it is mostly frequent among the Old Order Amish society(Slavotinek, 2017). Heart malformations have been defined at least once in patients with MKS including atrioventricularis (AV), atrial and ventricular septal defects, small aorta and hypoplastic LV, TOF, and PDA. The relative incidences are not clear because of the limited number of patients with MKS and heart malformations(Slavotinek & Biesecker, 2000).

D. LIM domain proteins

LIM domain proteins were described more than 20 years ago and are now known as highly conserved zinc-finger motifs(Freyd, Kim, & Horvitz, 1990). The LIM domain was first characterized in cDNA encoded by *Caenorhabditis elegans* specifically in cysteine-rich sequence *MEC-3* gene which is essential for the specification of the mechanosensory neurons. There were no comparable sequences present in databases at that time until the cloning LIN-11 and Isl1 genes, which encode *Caenorhabditis elegans* lineage protein and rat insulin gene-enhancer binding protein respectively(Kadrmas & Beckerle, 2004). Thus, the abbreviation "LIM" refers to first letters of three homeodomain transcription factors that are involved in regulation: a cell lineage protein Lin-1, Insulin enhancer-binding protein Isl1, and a mechanosensory neuron differentiation protein Mec3. This origin proposes that LIM domains play a role in gene expression, embryonic development and stress/strain detectors (A. Li, Ponten, & dos Remedios, 2012). The LIM proteins are also involved in other key roles such as cytoskeleton organization, cell differentiation, signal transduction, and cardiomyocyte metabolism. Dawid *et al.* divided LIM proteins into groups according to the position of the LIM motifs in their sequences (Dawid, Breen, & Toyama, 1998). LIM domain

proteins were then divided according to their sequence homologies into four broad categories (Zheng & Zhao, 2007) . The first group comprises the N-terminal tandem LIM domain proteins such as LHX and nuclear LIM Only proteins (LMO), which are localized in the nucleus and function as a transcription factors or cofactors. Second group of proteins consist of LIM-only proteins that can shuttle between nucleus and cytoplasm. These protein families include CRP, FHL and PINCH. The third group proteins usually contains a C-terminal LIM domain plus several other protein-protein interaction motifs such as PDZ, LD (leucine-aspartate repeat) and ATD (actin-target domain). These proteins include protein families Paxillin, Zyxin, Testin, Enigma, and some other proteins. The fourth group of proteins carries in addition to LIM and protein-protein interaction domains also mono-oxygenase or kinase catalytic motifs, like LIMK and MICAL proteins (Kadrmas & Beckerle, 2004) (figure 13).

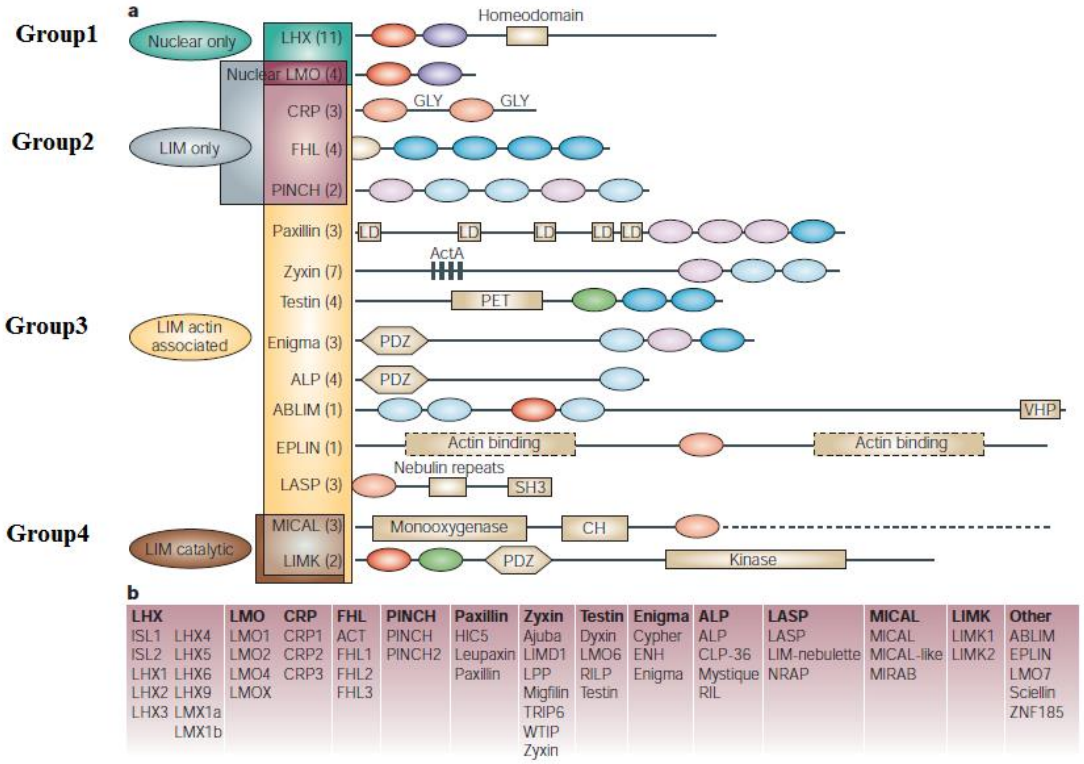


Figure 11. Human LIM-domain protein groups. The schematic drawing of domain structure of the founding member and/or best characterized example of the 4 groups of LIM-protein families are presented. Between parentheses, the number of the known family members is indicated. The most common categorization schemes are presented by colored boxes. The description of heterologous domains such as LD motifs, actin-binding domains and monooxygenase domains is mentioned in text. The dashes show the domains with boundaries that have not specifically been identified (a). The box shows recognized members of each LIM family (b).ABLIM, actin-binding LIM protein; ACT, activator of cyclic AMP response element modulator (CREM) in the testis; ALP, α-actinin-associated LIM protein; CH, calponin homology; CRP, cysteine-rich protein; EPLIN, epithelial protein lost in neoplasm; FHL, four-and-a-half LIM; GLY, glycine-rich region; LASP, LIM and SH3 protein; LHX, LIM-homeodomain protein; LIMK, LIM kinase; LMO, LIM only; MICAL, molecule interacting with CASL protein-1; PDZ, postsynaptic density-95, Discs large, zona occludens-1; PET, prickle, espinas and testin; PINCH, particularly interesting new cysteine and histidine-rich protein; SH3, Srchomology- 3; VHP, villin head piece(Adapted from Kadrmas & Beckerle, 2004)

1. LIM domain sequence and structure

The LIM domain contains a conserved double zinc finger motif that is evolutionary conserved and is found in a variety of proteins displaying distinct biological roles. LIM domains have been identified a wide range of eukaryotes in which 135 LIM domains have been identified within 58 genes. The LIM domains contain highly conserved residues, mainly the cysteine and Histidine residues(Kadrmas & Beckerle, 2004). All the LIM domains are made up of 50-60 amino acids and are characterized by the following sequence CX2CX16–23HX2CX2CX2CX16–21CX2(C/H/D) in which X refers to any amino acid (A. Li et al., 2012; Zheng & Zhao, 2007). Though, the LIM domain sequences are still diverse between species and splice variants aside from the core sequence of cysteine (C) and histidine (H) amino acids in which a larger LIM sequence has been detected in humans such as C (X) 2C (X)16–23 (H/C) (X) 2/4 (C/H/E) (X)2C (X)14–21(C/H)(X)2/1/3 (C/H/D/E)X (A. Li et al., 2012). Since zinc-finger motifs are closely similar to GATA-type zinc fingers, it was suggested

that they might bind specific sequences on DNA. It was found that LIM domains of ISL1 and MEC3 proteins hinders the binding of homeodomains to their DNA sequences (Sánchez-García, Osada, Forster, & Rabbitts, 1993; Xue, Tu, & Chalfie, 1993). Yet, there is no direct evidence that supports the binding of LIM domains to DNA even in LIM domain transcription factors(Zheng & Zhao, 2007). Interaction of LIM domains with specific protein partners is now known to influence its subcellular localization and activity(Khurana, Khurana, & Noegel, 2002).The conserved features present in the LIM domain help facilitate the formation of stable structures and different features that impart high-affinity binding to several structurally and functionally diverse protein partners(Kadrmas & Beckerle, 2004).

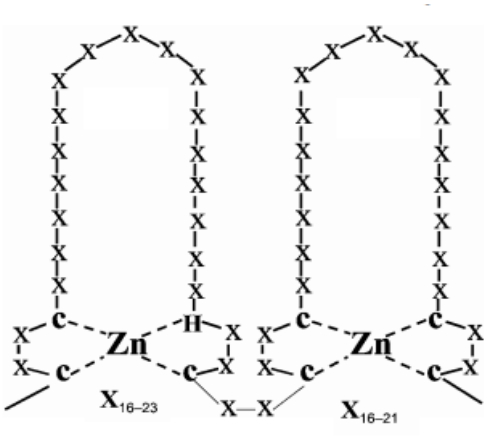


Figure 12. The structure of the LIM domain. The LIM domains are commonly made up of 50 to 60 amino acids in size and harbor two typical zinc fingers that are separated by two amino acids. The bold letters represent the conserved cysteine (C) and histidine (H) residues that make up the two zinc fingers (Adapted from Zheng & Zhao, 2007).

a. The Family of Cysteine Rich Proteins (CRP)

In humans, three CRP-family members (group two of LIM domain proteins) have been identified which are: CRP1, CRP2 and CRP3/MLP (Table 6). CRPs are small proteins, 22 kDa of size, and contain two functional LIM domains that are linked to glycine-rich repeat (figure 13)(Kadrmas & Beckerle, 2004).

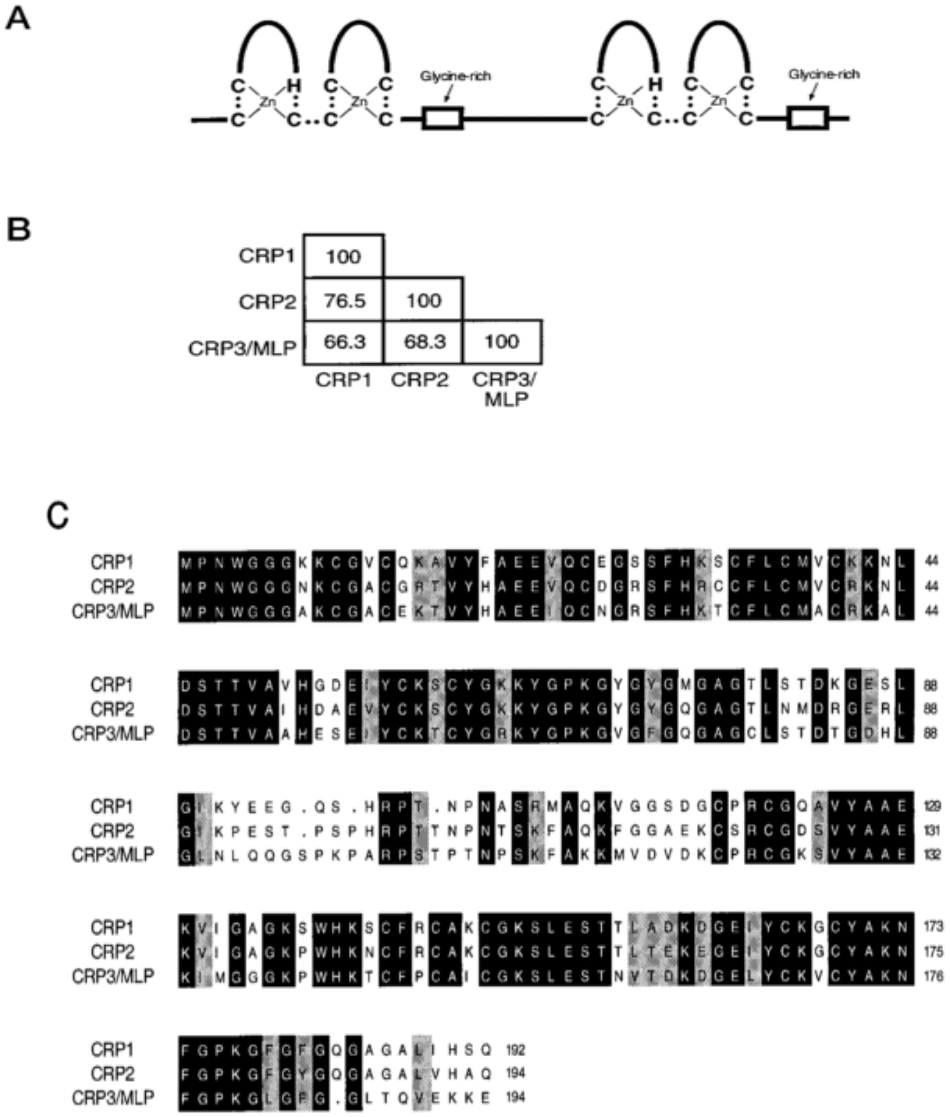


Figure 13. **The three highly conserved proteins (CRP1, CRP2, and CRP3) of the chicken CRP multigene family.** The molecular construction of CRP family members is characterized by the existence of two zinc-binding LIM domains which are followed by a conserved glycine-rich repeat (A). The identities and pairwise comparisons between chicken CRP sequences done by Lipman and Pearson reveal that that CRP1(Lipman & Pearson, 1985), CRP2, and CRP3/ MLP have high percentage of sequence identity (B). The alignment of all the three chicken CRP amino acid sequences demonstrates the specific regions within these proteins that have the highest amount of similarity. Black boxes= residues that are totally identical between the 3 protein molecules, gray boxes= residues that, while not identical between all three proteins, exhibit prominent sequence similarity (As previously described, the similarity groups were assigned(R. Weiskirchen, Pino, Macalma, Bister, & Beckerle, 1995): Ala, Ser, Thr; Asp, Glu; Asn, Gln; Arg, Lys; Ile, Leu, Met, Val; Phe, Tyr, Trp). The sequence of the chicken CRP3/MLP shown here varies in four positions from the previously described sequence (at amino acids 92, 93, 94, and 114)(Arber, Halder, & Caroni, 1994).The corrections to the CRP3/MLP sequence that are described here result in slightly larger degrees of similarity between CRP isoforms than were reported previously (C)(Arber et al., 1994; R. Weiskirchen et al., 1995) (adapted from Louis et al. 1997).

Several related proteins have been found to share common features with the CRP vertebrate counterparts, suggesting that these proteins are evolutionarily conserved. CRP like LIM domain proteins have been known in plants, arthropods, and protozoas. In *Drosophila melanogaster*, two members of CRP counterparts have been found and referred to as MLP(Ralf Weiskirchen & Günther, 2003) that share features with vertebrate isoforms, Mlp60A and Mlp84B. Although both the isoforms are co-expressed in somatic, visceral, and pharyngeal muscles, they exhibit distinct subcellular distributions. Whereas Mlp60A is distributed throughout the muscle fibers, Mlp84B is distributed to the sites of muscle attachment and the periphery of Z-bands of striated muscle, which is rich in α-actinin(Stronach, Siegrist, & Beckerle, 1996).Moreover, their expression is enhanced during development, supporting their structural role in late myogenesis(Stronach et al., 1996) . CRP counterpart in *Dictyostelium discoideum* is known as DsLIM. Although it has similarities to CRP proteins, DdLIM contains one LIM domain only. However, it has no putative nuclear targeting signal and a very diffuse glycine-rich repeat (Prassler et al., 1998).

CRP family members play a role in terminal differentiation in vertebrate muscle development. CRP1 and CRP2 are prominent in smooth muscle; and CRP3, which is also known as the muscle LIM protein (MLP), is expressed in striated muscle. In the cytoplasm, all three proteins interact with α-actinin (an actin cross-linking protein) and are associated with the actin cytoskeleton (Louis et al., 1997). They were also shown to interact with zyxin which is the adhesion plaque LIM domain protein (Louis et al., 1997; Pomiès, Louis, & Beckerle, 1997; Sadler, Crawford, Michelsen, & Beckerle, 1992; Schmeichel & Beckerle, 1994). The binding of zyxin to CRP1 appears to rely on

the coordinate action of sequences found in both of CRP1's LIM domains (Schmeichel & Beckerle, 1998).

In chicken, CRPs have a variable expression pattern. For example, CRP1 is expressed in most of the tissues and is specifically highly expressed in muscle cells. However, CRP2 expression is likely to be restricted to arteries and fibroblasts while CRP3/MLP is in striated muscles(Louis et al., 1997). Though, when CRP proteins are expressed in the fibroblasts, they share a common feature in which they all are associated to actin cytoskeleton and interact with α-actinin and zyxin(Louis et al., 1997). It is predicted that LIM2 which is located at the C-terminal of CRPs might have more than one partner. For example, CRP3/MLP also interacts with β1-spectrin(Flick & Konieczny, 2000). Consequently, although CRPs have common interacting partners, it is predicted that binding proteins could be unique to each CRP. Thus, this allows different functions of CRPs in the cell(Järvinen, 2011).

b. Cysteine Rich Protein 1 (CSRP1/ CRP1)

In 1990, Liebher et al. were the first to clone the human CRP1 coding gene (*CSRP1*)(Liebhaber et al. , 1990). The cDNA length was estimated to be 1778 bp, and the gene harbors double domains each consists of two putative zinc finger domains in addition to glycine-rich repeats. From yeast to humans, CRP gene is highly conserved throughout evolution. In human genome, it is present as single copy and localizes to chromosome 1q24-1q32(Erdel & Weiskirchen, 1998; Xinkang Wang, Leesv, Liebhabersllll, & Cookes, 1992). CRP1 was also cloned in chicken and the three dimensional structure was studied by NMR; multidimensional nuclear magnetic

resonance (Crawford, Pino, & Beckerle, 1994; Yao et al., 1999). The latter revealed that LIM domains do not interact and are spatially separated which suggested that they might be acting independently as linkers or adaptors (figure 14). A putative nuclear localization signal (NLS) was located in CRP1 sequence(R. Weiskirchen et al., 1995). Studies have shown that there are some structural similarities between CRP1 and CRIP (cysteine-rich intestinal protein). While CRIP contains two zinc fingers solely, CRP1 consists of four zinc fingers containing a 25-amino acid motif. Since there is a sequence similarity between zinc fingers 1-3 and 2-4, CRP1 is suggested to have evolved as a duplication of CRIP gene (Liebhaber et al., 1990). Four CRP1 transcripts have been detected. Since CRP1 and CRP2 gene have highly identical sequences, it is predicted that CRP genes have probably been duplicated from a common gene ancestor (Brenda Lilly, Olson, & Beckerle, 2001).

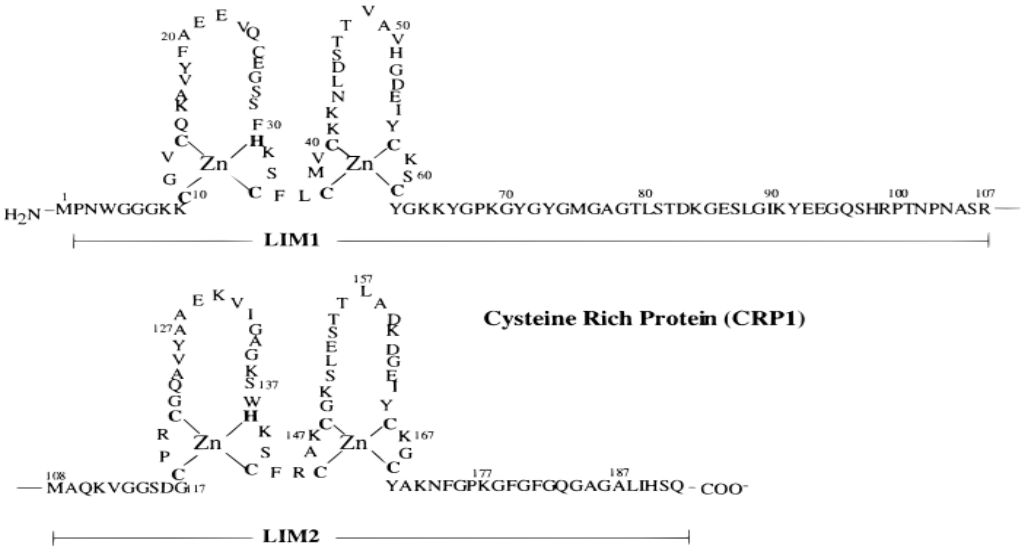


Figure 14. LIM1 and LIM2 zinc-binding domains. This diagram shows the metal-binding mode and the amino acid sequence of Cysteine rich protein 1 (CRP1) containing two tandemly arranged LIM1 and LIM2 domains respectively(Adapted from Yao et al., 1999).

Table 6. The expression of CRP proteins, their binding partners and their functions (adapted from Järvinen, 2011).

PROTEIN	EXPRESSION IN ADULT TISSUES	BINDING PROTEINS	KNOCK-OUT/KNOCK-IN PHENOTYPE
CRP1	in vascular and avascular tissues containing smooth muscle cells	α-actinin, Zyxin, actin, SRF, GATA6, PKC, Dishevelled, Diversin	Knock-out mice: Viable and Fertile Attenuated neoitima formation after balloon surgery No effect on smooth muscle differentiation knock-down in *zebrafish:* abnormal cell movement during development
CRP2	Mainly in smooth muscle cells of vasculature	α-actinin, Zyxin, actin, STAT1, SRF, GATA6,CRP2PB	Knock-out in mice: Viable and fertile No effect on smooth muscle differentiation No effect of vascular Enhanced vascular smooth muscle migration Subtle changes in cardiac Ultrastructure Knock-in mice: transgenic CRP2 switches on smooth muscle gene expression in cardiomyocytes
CRP3/MLP	In cardiac and Skeletal muscles	α-actinin, Zyxin, cofilin, MyoD, myogenin, MRF4,β-spectrin	Knock-out mice: viable and born in expected ratios hearts abnormally soft, cardiomyocytes show disruption of cyto-architecture, leads to the development of dilated cardiomyopathy Mutations found in families with hypertrophic cardiomyopathy

i. Regulation of CSRP1 expression

CSRP1 has been suggested to function as a smooth muscle markers since it is expressed in smooth muscle containing vascular and nonvascular tissues (Henderson et al., 1999; Yet et al., 1998). Chicken CSRP1 protein is highly expressed in gizzard, intestine and stomach tissue which contain rich amount of smooth muscle cells(Crawford et al., 1994). In addition, northern blotting analysis revealed that Csrp1 mRNA in mice expression is high in the lung, kidney, and brain and is weaker in spleen, skeletal muscles, spleen, testis and heart (Henderson et al., 1999) (figure 17). In various mouse tissues, CSRP1 expression was prominently detected under more careful analysis. By using a hybridization probe, *Csrp1* transcript was not detected in cardiac tissue itself, but was specifically found in smooth muscle cells of adult murine cardiac

arteries . In addition, CRP1 did not show any significant expression in the secretory epithelium of human prostate, however, it was highly expressed in the cytoplasm of stromal compartment (Dube et al., 1998).

Ca^{2+}/calmodulin (CaM)-dependent protein kinases consist of three members including CaMKI, CaMKII and CaMKIV. CRP1 expression was induced by CaMKIV but not with CaMKII by CRE and CArG element regions in the promoter(Najwer & Lilly, 2005). Serum response factor (SRF) controls the transcription of numerous smooth muscle specific genes that contain CArG box enhancer elements which leads to the organization of actin cytoskeleton in response to various stimuli (Olson & Nordheim, 2010). This regulation is mediated through the binding of SRF to CArG element on the promoter. The CArG-element in Csrp1 is situated in the 5 kb enhancer. SRF binding to this element is essential to regulate Csrp1 expression in smooth muscles of the arteries but not in veins or viscera (Brenda Lilly et al., 2001).

ii. Role and Expression of CSRP Proteins in the Heart

In Humans, hundreds of LIM domain proteins have been defined (Consortium, 2011), however, only few have been involved in cardiac development and/ or dysfunction including the cysteine rich proteins (figure 15).

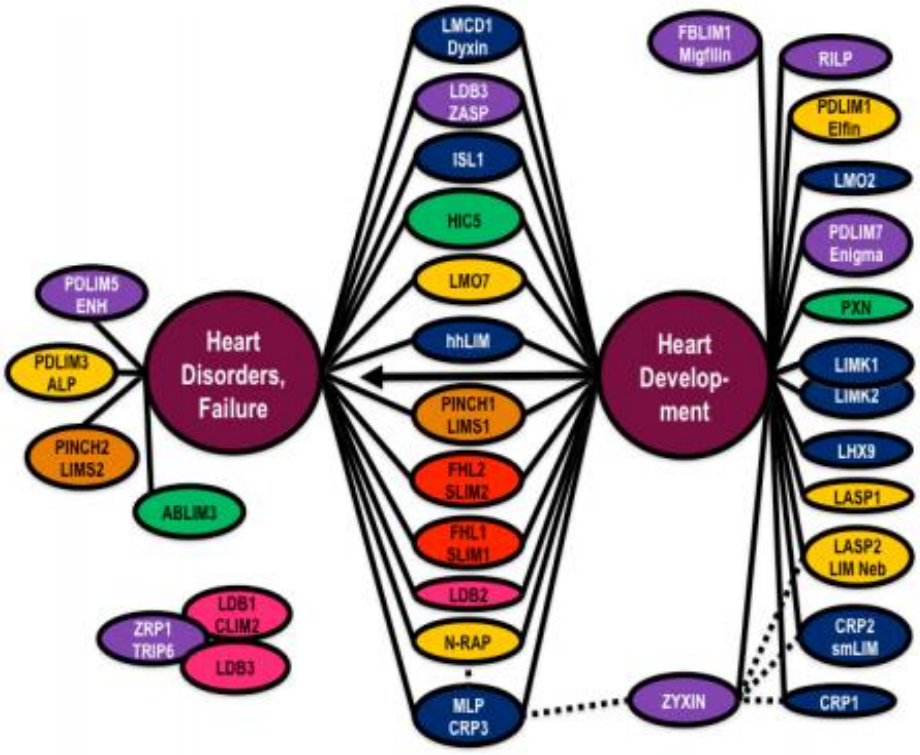


Figure 15. Cardiac and blood vessel LIM domain proteins. Those are arranged according to their functional contribution to heart development and/or disorders (and failure). LIM proteins on the right have been linked to embryonic or post-natal heart development only. However, the ones on the left are only associated with cardiac disorders or cardiac failure. The LIM proteins in the middle are associated with both heart development and disorders and in some situations the transition to heart failure. TRIP6, LDB1 and LDB3 are orphan proteins since they are expressed in the heart but they have unknown function. Dashed lines designate interactions between the LIM domain proteins (A. Li et al., 2012).

During embryonic development, CRP1 participates in the formation of the cardiac mesoderm and its down regulation alters mesoderm cell migration patterns resulting in *cardiac Bifida* (the presence of totally or partially separated hearts). In adult mouse, CRP1 is expressed in muscle cells of cardiac arteries(Y. Q. Yang, Li, Wang, Liu, Chen, Zhang, Wang, Bhuiyan, et al., 2012). A study done on zebrafish model showed that CRP1 acts as an important factor during gastrulation and cell movement of the mesoderm and cardiac mesoderm. Wnt signalling pathway is essential for convergent extension during morphogenesis. CRP1 was shown to interact with Wnt signalling pathway components Dishevelled and Diversin. CRP1 inhibition leads to

irregular cell movement in convergent extension resulting in anomalies in midline structures(Miyasaka, Kida, Sato, Minami, & Ogura, 2007) . LIM domains of different LIM domain proteins have also been noticed to interact with PKC, as was the case with CRP1(Kuroda et al., 1996). However, CRP1 does not affect the activation of PKC(Maturana et al., 2011) .

CRP2 was found to be transiently expressed in heart during embryogenesis and to have a role in cardiac muscle differentiation. Cardiomyocyte-specific expression of transgenic CRP2 switches on smooth muscle gene expression in cardiac myocytes in mice. CRP2 was suggested to function as a transcriptional co-adaptor protein(Chang, Belaguli, Chang, & Schwartz, 2007). Moreover, *Csrp2* gene knock-out resulted in obvious changes in cardiac ultrastructure, although mice with non-functional CRP2 were otherwise viable and fertile. Under such condition, the thickness of the cardiomyocytes was increased and the cells were hypertrophic(Sagave et al., 2008).

CRP3/MLP which was first isolated as a cDNA from rat skeletal muscle(Arber et al., 1994) is a positive regulator of myogenic differentiation. The overexpression of CRP3 in myoblasts enhances skeletal myogenesis. On the other hand, *Csrp3* silencing in myoblasts prevents them from exiting cell cycle and block terminal differentiation. At the beginning of muscle differentiation, CRP3 accumulates into nuclei and is observed in the cytoplasm later during development(Arber et al., 1994). Furthermore, CRP3 has been identified as a cofactor in myogenesis inducing complex, which interacts in nucleus with Myogenin, MyoD, and MRF4 through its first LIM domain and the helix-loop-helix motifs of the MRFs(Kong, Flick, Kudla, & Konieczny, 1997).

iii. The dual expression of CSRP proteins in the cell

Many LIM proteins that were initially identified as cytoskeleton-associated proteins, such as members of the cysteine-rich protein (CRP) families, four-and-a-half LIM (FHL), PINCH and zyxin are recognized to shuttle between the cytoplasm and nucleus of the cell to influence gene expression(Cattaruzza, Lattrich, & Hecker, 2004; Chang et al., 2003b). This dual localization is due to the presence of a putative nuclear targeting signal (KKYGPK) that has been identified in the glycine-rich regions of the CRPs(Arber & Caroni, 1996; Arber et al., 1994; Stronach et al., 1996) (figure 16).

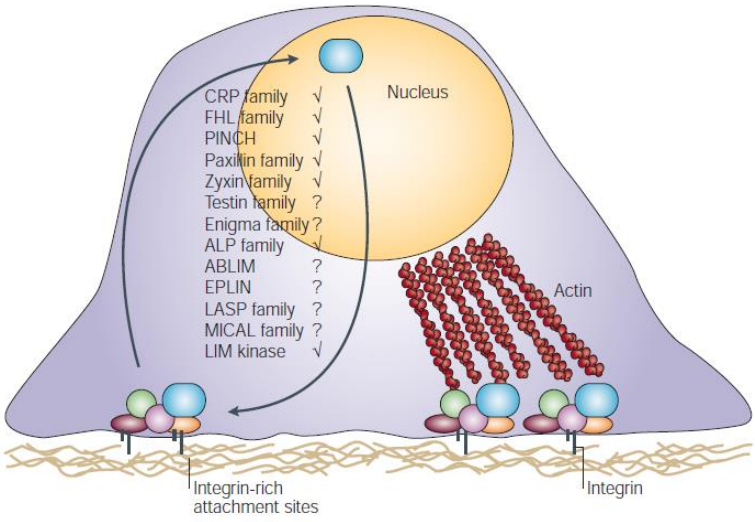


Figure 16. The dual localization of LIM domain proteins. A list of LIM proteins that have been observed at focal adhesions, muscle-attachment sites or other analogous integrin-rich attachment structures is shown in the figure. Some of these LIM domain proteins can also be localized in the nucleus (√), even though many of these proteins that belong to this group have no examined nuclear localization signal (?). This dual localization might be need to allow communication between these cellular partitions. Actin-binding LIM protein: ABLIM, α-actinin-associated LIM protein: ALP; cysteine-rich protein: CRP; epithelial protein lost in neoplasm: EPLIN; four-and-a-half LIM: FHL; LIM and Src-homology-3 protein: LASP; molecule interacting with CASL: MICAL; particularly interesting new cysteine and histidine-rich protein: PINCH (Kadrmas & Beckerle, 2004).

The LIM domain proteins usually function in tissue-specific regulation and cell fate determination in the nucleus. However, LIM domain proteins localized in the

cytoplasm play a role in cytoskeleton organization(Zheng & Zhao, 2007). In addition, other islet-expressing tissues, such as heart and alary muscles, exhibit *ISL1* gene. The nuclear transcription factor ISL1 with its target (MEF2C) is involved in the development of the right ventricle and cardiac outflow tracts which form the second heart field. Mutations in *Isl1* are involved in congenital heart defects (e.g. Tetralogy of Fallot, transposition of great vessels, tricuspid atresia)(Giuliano, Marino, Pinto, & De Santis, 1998). This proposes the importance of this gene in driving cardiac progenitors' differentiation in the second heart field. A study done by Laugwitz *et al.* have showed that ISL1 progenitors are noticed at lower levels in the post-natal primary heart field myocardium (interatrial septum, left atrium and LV) in humans , mice, and rats. It was further identified that subsets of ISL1 progenitor-derived cardiomyocytes showed conduction, atrial and ventricular cell electrophysiological characteristics(Pfaff, Mendelsohn, Stewart, Edlund, & Jessell, 1996).

It is clear now that CRP-family members enter to the nucleus to promote muscle-specific gene expression(Arber & Caroni, 1996; Chang et al., 2003b). It has been obviously established how the CRPs contribute to the modulation of gene expression. CRP1 and CRP2 were shown to function through coordinated docking of Serum-Response Factor (SRF) to the N-terminal LIM domain and GATA factors, specifically GATA4 and GATA6, to the C-terminal LIM domain (Chang et al., 2003b). The strong expression of many smooth muscle- differentiation markers is stimulated by this ternary complex of SRF–CRP–GATA, whereas the pairwise combinations have much less impact on gene expression. It is tempting to speculate that CRPs monitor the integrity of the muscle contractile machinery and contribute to its homeostasis, given the dual subcellular distribution of CRP and its clear transcriptional role. If this is the

case, CRPs that are associated with the actin cytoskeleton might function as sensors to assess the physiological status of the contractile machinery(Kadrmas & Beckerle, 2004).

When muscle is damaged or placed under stress, the CRPs might signal to the nucleus in order to signal for repair through the activation of muscle-specific genes. This hypothesis is supported by the fact that CRP3/MLP seems to be a component of a Z-disc-localized stress sensor in cardiomyocytes and leads to the development of human dilated cardiomyopathy when it is defective(Knöll et al., 2002). The cardiomyocytes respond to mechanical stress through alterations in gene expression that lead to hypertrophy, and CRP3/MLP82 shows a nuclear relocation in rats with cardiac hypertrophy(Ecarnot-Laubriet et al., 2000). The CRP family demonstrates not only the way in which multi-LIM proteins can function as a framework for the generation of multi-component regulatory machines, but also the theme of LIM proteins as connectors between the cytoskeleton and the nucleus(Kadrmas & Beckerle, 2004).

E. Significance of our study

It has been previously reported that CSRP1 is expressed during cardiovascular development and acts as a multifunctional protein in the specification of cardiovascular lineages (Chang et al., 2003a). Previous studies have shown that CSRP1 is highly expressed in regions enriched with smooth muscles including arteries in adults(Henderson et al., 1999). However, no previous study has shown its expression in mammalian embryonic or adult heart. Although several studies have associated the role of CSRP1 with embryonic heart development in *zebrafish* and *chick* embryos, no previous study has related a mutation in *CSRP1* to CHD. This study was a challenge

since our targeted sequencing data on a set of genes from a Lebanese family having various forms of CHD in addition to polydactyly revealed a novel heterozygote mutation in *CSRP1*. This will help in deciphering the mysteries of heart formation and would lead to new approaches that could be used to prevent the death of neonates with CHD.

CHAPTER II

AIMS OF THE PROJECT

Abnormal cardiac development leads to human congenital heart disease (CHD), which is responsible for the vast majority of neonates' death around the world. DHFMR-85 is a big Lebanese family composed from the consanguineous marriage between two first-degree cousins. Out of the seven conceived children, two died *in utero* at the ages of six and nine months of unknown causes. Of the remaining five children, three have congenital heart disease (ventricular septal defect, atrial septal defect, and patent ductus arteriosus), and four have polydactyly. Targeted exome sequencing identified a heterozygote duplication of a 14-nucleotide fragment in exon 5 of *CSRP1*, causing a frameshift mutation at position 154 of the protein. CSRP1 belongs to the LIM-only proteins family which associate with different partners in the cytoplasm and the nucleus and therefore are involved in various cellular functions including cytoskeleton re-modeling versus regulation of gene expression. Previous studies done on *zebrafish* have shown that CSRP1 participates in the formation of cardiac mesoderm and its down regulation resulted in *cardiac bifida*. In adult mouse, CSRP1 was shown to be expressed in smooth muscle cells of cardiovascular system. Mutations of a small but growing number of genes have been shown to cause CHD. No previous study has linked a mutation in *CSRP1* to congenital heart disease or cardiac development.

Therefore, the objective of our project is to show that the novel frameshift mutation, p.E154Vfs*99, detected in CSRP1 is disease-causing. By this, we suggest CSRP1 as a potential biomarker for CHD.

The specific aims of the work are:

A. Specific Aim 1: *In silico* analysis of CSRP1 mutation. To carry on this aim we will:

- Retrieve the coding sequence of *CSRP1* gene and the wild type protein by using bioinformatics tools. This will allow us to locate the conserved LIM domains on CSRP1 gene and protein.
- Align the WT and MUT CSRP1 protein sequences. This will help us assess how the mutation is affecting the protein sequence and to map which domain is affected by the mutation.

B. Specific Aim 2: Characterize the *in vivo* expression of CSRP1 in cardiac tissues. To carry on this aim we will:

- Examine the expression and localization of CSRP1 in mouse cardiac tissues at different developmental stages by immunohistochemistry. The expression of CSRP1 in cardiomyocyte will confirm its potential role in cardiac development.
- Examine the expression of CSRP1 at the septal region in mouse cardiac tissues during development by immunohistochemistry. The expression of CSRP1 in the septal region will reveal a possible role for the protein in septal formation.

C. Specific Aim 3: Characterize the function of WT and MUT (p.E154Vfs*99) CSRP1 *in vitro*. To do this we will:

- Construct the WT *CSRP1* cDNA and design the MUT *CSRP1* by site-directed mutagenesis and then check protein expression of WT and MUT CSRP1 by immunoblotting after transfecting the WT and MUT CSRP1 in HEK293 cells.

- Examine the effect of p.E154Vfs*99 mutation on the cellular localization of CSRP1 protein by immunofluorescence after expressing the WT and MUT CSRP1 in Hela cells through transfection.
- Assess the transcriptional activity of WT and MUT CSRP1 on cardiac developmental promoters such as *NPPA*, *VEGF* and *NOS3* through transfections in HEK293 cells and by performing luciferase assay.
- Determine whether there is a physical interaction between CSRP1 and TBX5. This will be done by performing co-immunoprecipitation. We will determine whether the mutation is affecting the physical interaction between CSRP1 and TBX5.
- Determine whether there is a physical association between CSRP1 and nuclear proteins GATA transcription factors including GATA4, 5 and 6 in addition to SRF and whether this interaction is abrogated by p.E154Vfs*99 CSRP1 by co-immunoprecipitation.
- Examine if the functional interaction determined between WT CSRP1 and GATA transcription factors and/or SRF is disrupted by p.E154Vfs*99 mutation through performing luciferase assays on HEK293 cells.
- Identify other genes that might explain the cardiac phenotype of the family members. This part will be fulfilled by whole-exome sequencing and subsequent *in vitro* assays.

CHAPTER III

MATERIALS AND METHODS

A. Patient Recruitment and clinical examination

The study was approved by the institutional review board at the American University of Beirut (protocol number: Bioch.GN.01). All patients, and family members signed an informed consent form before being involved in the study. A total of 20 individuals from the same family were enrolled. Standard clinical evaluation included a complete physical exam, electrocardiography (ECG), and two-dimensional (2D) transthoracic echocardiography (TTE) with color Doppler was attained. After interviewing all patients and parents, family consanguinity history was utilized in constructing the pedigree.

B. Genetic analysis

Peripheral venous blood was collected from each member. Genomic DNA was extracted from white blood cells using the Qiagen Blood-Midi kit (Qiagen Science Inc., Germantown, MD), as previously described (38). Primers to amplify all coding exons were designed using genome.ucsc.edu PCR design. Amplification by polymerase chain reaction (PCR) was done using the Phusion polymerase high-fidelity master mix (F-548S) on a Pico machine (Finnzymes, Espo, Finland), and the amplicons were resolved on a 1.5% agarose gel. Gel purification was performed using the Gel Extraction kit following the manufacturer's protocol (GenElute^{TM} Gel Extraction Kit: Catalog Number NA1111, Sigma-Aldrich). After purification of bands, DNA was quantified using a

Nano Drop (DeNovix DS-11) and then examined by gel electrophoresis in order to guarantee quality.

C. DNA Sanger-Sequencing

DNA sequencing was carried out on an ABI 3500 machine at the molecular core facility at the American University of Beirut, followed by analysis using the data collection software from Applied Biosystems Inc. (Foster City, CA).

D. Cell lines

HEK293 cells (Human Embryonic Kidney cells), Hela cells (human cervical cancer cells) were cultured and maintained in Dulbecco's Modified Eagle Media (DMEM- Sigma, Cat#D0819) supplemented with 10% Fetal Bovine Serum (FBS-Sigma, Cat#F9665), 1% Penicillin/ Streptomycin (Biowest-Cat#L0022-100) and 1% Sodium pyruvate (Sigma-Cat#S8636). Incubation of cells was carried out in a 5% CO2 humid atmosphere at 37ºC.

E. Site-directed Mutagenesis

1. Primer phosphorylation

The primers for the p.E154Vfs*99 CSRP1 and p.R311S TRPS1 (Table 10) were first phosphorylated according to the following protocol. Forward and reverse primers were diluted 1:20 and then 1 µl of each primer was incubated along with 1 µl T4 polynucleotide kinase (PNK) (Thermo scientific, Cat# EK0031) , 5 µl ATP (10 mM), 2 µl (1/10 of total sample volume) Thermo Scientific 10x reaction buffer A , and the volume was completed to 20 µl by adding nuclease-free water. The samples were mixed

thoroughly, centrifuged briefly and then incubated at 37°C for 30 minutes after that at -20°C for 20 minutes.

2. Polymerase Chain Reaction (PCR)

The mutated site was amplified in a 25 µl reaction mixture using 2X Phusion polymerase Kit (Thermo Scientific) Cat# F531S under the following conditions: 3 µl of forward primer, 3 µl of reverse primer, 10 ng of the plasmid with the mutated site, 12.5 µl of Phusion enzyme and the volume was completed to 25 µl with nuclease free water. PCR conditions on T100 Thermal Cycler (BIORAD) were : 98ºC for 60 seconds as an initial denaturing step; 98ºC for 20 seconds, 58ºC for 20 seconds, 72ºC for 90 seconds for 30 cycles; a final extension step at 72ºC for 10 minutes followed by a hold temperature at 4ºC.

3. Gel electrophoresis

PCR products were loaded on 1.5% agarose gel with a loading dye (0.25 % Bromophenol blue, 0.25 % Xylene Cyanol and 15 % glycerol) in a 6 to 1 ratio, along with PSK/ HpaII molecular ladder (20 µl plasmid with 3 µl HpaII, 5 µl one phor all (OPA) buffer and 22 µl sterile water). Agarose gel was prepared in TBE 1X (Tris Boric EDTA, 1L of 10X contain 108g Tris base, 55g Boric acid and 40ml EDTA 0.5 M pH 8) and stained with 1 mg/ml ethidium bromide to visualize nucleic acids.

4. PCR product purification

PCR products were purified from TBE agarose gel band using Nucleospin extraction kit Cat# 740609.10 following the manufacturer's protocol.

5. Ligation

To prepare the ligation mixture, the purified inserts and plasmids were added in 3:1 ratio, 1 µl T4 DNA ligase (Thermo Scientific) Cat# K1239, 2 µl of 10x T4 DNA ligation buffer Lot#00147820 and water up to 20 µl. Negative controls were prepared following the same procedure where no insert was added. Then, the ligation mixture was incubated at 37°C for 30 minutes and then at room temperature for 2 hours.

6. Transformation in bacteria

The previously obtained constructs were then transformed into *E.coli, XL1 blue bacteria* strain stored at -80 ºC. In eppendorf tubes, 100 µl of bacteria is added to 1-2 µg of the plasmids containing our DNA constructs. The eppendorf tube is then inverted up and down for several times, placed 5 minutes on ice, 5 minutes at 37ºC (in the water bath), then 2 minutes on ice.

The transformed bacteria are streaked on agar plates, and then incubated overnight at 37ºC. It is important to mention that this transformation process is performed in aseptic technique i.e. close to the flame of a Bunsen burner. Since the undigested plasmid contains an ampicillin-resistance gene, and since during the preparation of the agar plates, ampicillin is added, only bacteria that took up the ligated plasmid with the insert can grow. Negative control should contain no colonies since the digested plasmid should not ligate to itself. Bacterial colonies observed to grow on the agar are removed with pipette tips, and are then transferred into 15-mL falcon tubes containing 3 mL liquid broth with 3 µl ampicillin (100 mg/ml). The tubes are then incubated overnight in the shaker, at 37ºC, at 150 rpm. Miniprep and maxiprep are performed using illustra™ plasmidPrep Midi Flow Kit (GE Healthcare) according to the

enclosed manufacturer's protocol. Sequenced plasmids containing our required constructs are purified.

F. Expression Vectors and Reporter Genes

Luciferase reporter plasmids PGL3-VEGF-luc, PGL2-NOS3-luc, and PXP2-NPPA-luc were constructed by ligation of PCR –amplified fragments from mouse *VEGF*, human *NOS3* and human *NPPA* promoters into eukaryotic luciferase expression vectors PGL3, PGL2 and pXP2 respectively. Flag-tagged wild type human CSRP1 was generated by sub cloning CSRP1 fragments into expression vector pCEP4 (Invitrogen). Flag-tagged p.E154Vfs*99 CSRP1 was constructed by ligation of PCR amplified fragments into eukaryotic expression vector pCEP4. PCGN-HA-SRF was a generous gift from Dr. Mona Nemer, Ottawa, Canada. HA-tagged pCGN-GATA (4, 5, and 6) and pCGN-TBX5 were cloned into eukaryotic expression vector pCGN. pCMV3-HA-TRSP1 was from Sino Biological Inc. (Cat# HG15989-NY).

G. Antibodies

Rabbit polyclonal antibody to CSRP1 Cat#ab70010; Goat anti-mouse antibodies (HRP) Cat# ab6789; and goat anti-rabbit antibodies (HRP) Cat# ab97051 were from abcam. Mouse monoclonal antibody against flag-tag (OCTA-probe H-5) Cat# sc-166355and rabbit polyclonal antibodies against HA-tag (HA-probe Y-11) Cat# sc-805 were from Santa Cruz. Mouse Biotinylated species-specific whole antibody (from donkey), Cat#: LRPN1001V and rabbit biotinylated species-specific whole antibody (from sheep); Cat# RPN1004V were from GE healthcare UK limited. Streptavidin Red full length Cat#ab136227 was from abcam. Chromeo ^{TM} 488 streptavidin green; Cat#

sc-364698 was from Santa Cruz. Goat anti-rabbit Alexa fluor 488 was from Invitrogen. Rabbit Ig-G, Polyclonal isotype Control Cat# abcam 27478 used to detect specificity of anti-CSRP1 was from abcam.

H. Transfection Assays to Assess CSRP1 Target gene promoters

HEK293 cells were grown and maintained at sub-confluence ~80% level in Dulbecco's modified Eagle's Medium (DMEM) containing 10% fetal bovine serum. Transient transfections were performed with Polyethylenimine (Sigma). A series of Luciferase assays were performed by transient transfection in combination with pCEP4-Flag-WT CSRP1, pCEP4-Flag-E154Vfs*99 CSRP1, PCGN-HA-SRF, pCGN-HA-GATA4, pCGN-HA-GATA4, pCMV3-HA-WT TRPS, luciferase reporters (3.5 µg) and empty expression vectors, PCGN, to a balanced total of 1 µg of plasmids per 2 wells of 12-well plate. The results were normalized to total protein concentration in each well, and were expressed as fold activation. Co-transfection experiments were performed in duplicates and repeated at least three times. Luciferase activity was normalized to baseline reporter gene activity as fold activation, with error bars representing SEM.

I. Protein Overexpression

HEK293 cells were transiently transfected with epitope-tagged vectors pCGN-HA-GATA4, pCEP4-Flag-WT CSRP1, pCEP4-Flag- p.E154Vfs*99 CSRP1, pCEP4-HA-TBX5, pCGN-HA-GATA5, pCGN-HA-GATA6, PCGN-HA-SRF, pCMV3-HA-WT TRSP1, and pCMV3-HA-p.R311S TRSP1 using Polyethylenimine (PEI,Sigma). HEK293 cells were plated in 100 mm corning culture plates until sub confluence ~80% determined by green fluorescent protein (GFP) transfection assay. After 24 hr, 20 µg of DNA and 35 µl PEI transfection reagent were added to an Eppendorf tube holding a

total volume of 1 ml DMEM medium. The mixture was vortexed for 10 sec. Then, the mixture was incubated 20 min at RT, and applied over the cells. Culture medium was changed after 3 hrs of transfection.

J. Nuclear protein extraction

Nuclear extracts from transfected HEK293 cells were obtained as per the following protocol. The cells were washed with 1 ml PBS (1X) two times. Then, 2 ml of PBS (1X) were added to the cell plates along with 50 µl EDTA (0.5 M; PH8), a chelating agent to detach the cells by placing then on a shaker for 20 mins. The cells were then collected into Eppendorf tubes and centrifuged for 90 sec at 11000 rpm (Fixed-angle rotor). The supernatant was then discarded and the pellet is suspended in 800 µl buffer A (10 mM Tris PH7.9, 10 mM KCl, 0.1 mM EDTA, 0.5 mM EGTA). Each 1 ml of buffer A was mixed with 1 µl DTT (0.5M reducing agent), 1 ul PMSF (0.5M) and 2 µl protease inhibitor cocktail tablets (Complete, Roche Diagnostics GmbH). The contents are mixed by inversion, and then the tubes were placed on ice for 15 min.

Then, 50 µL of NP40 10% per tube is added, vortexed, and then centrifuged for 90 seconds at 11000 rpm. A transparent pellet is obtained. Then, the supernatant (cytoplasmic extracts) is cautiously discarded, and the pellet is re-suspended in 100-200 µL buffer C (20mM tris PH7.9, 400 mM NaCl, 1 mM EDTA, 1 mM EGTA). Each 1 ml of buffer C is mixed with 1 µl DTT (0.5M reducing agent), 1 ul PMSF (0.5M) and 2 µl protease inhibitor (cocktail). The tubes are then placed on the shaker in a cold room (T=4ºC) for 20 minutes, after which the tubes are centrifuged for 90 sec at 11000 rpm,

the pellet (DNA) is discarded, and 30 μL aliquots of the obtained supernatant (nuclear proteins) are prepared, to be stored at -80ºC for future use.

K. Protein Quantification

Protein concentration of the of the nuclear protein extracts previously obtained was measured using the Bio-Rad DC protein assay kit, according to the manufacturer's protocol. The obtained concentrations are to be considered in further protein uses.

L. Western Blot

Equal amounts of nuclear cell extracts (20 μg protein) were re-suspended in 5X laemmlli buffer (1mL glycerol, 0.5mL βME, 3mL 10% SDS, 1.25mL 1M Tris pH6.7 and 2mg bromophenol blue). The samples were then boiled for 7 minutes and run on a denaturing SDS-PAGE for 1.5 hours. Then, they were transferred to a PVDF membrane (Amersham, GE healthcare) Cat#10600023. The membrane was blocked for 45 minutes at RT with shaking in TBT 5% skimmed dry milk (2.5g milk in 50 ml TBT). The blocked membrane was then incubated with the primary antibody (anti-flag against flag-tagged proteins or anti-HA against HA-tagged proteins) over night with shaking at 4 °C. The primary antibodies were diluted 1:1000 in 1% skimmed dry milk. After that, the membrane was incubated 1 hour at RT with secondary antibodies (anti-rabbit or anti-mouse) conjugated to horseradish-peroxidase (diluted 1/20000). Development was done using ECL ™ Western Blotting Detection Reagents (Amersham, GE healthcare) Cat# RPN2106. The protein bands were visualized by Chemidoc MP imaging system-Biorad and quantified using Image J software.

M. Coimmunoprecipitation (Co-IP)

After detecting WT CSRP1, p.E154Vfs*99 CSRP1, GATA4,-5,-6, SRF and TBX5 proteins by western blot, co-immunoprecipitation assay was done to assess the physical interaction between WT /p.E154V*99 CSRP1 (Flag-tagged), GATA4-6 (HA-tagged), SRF (HA-tagged), TBX5 (HA-tagged) and pCMV3-HA-TRSP1/ pCMV3-HA-p.R311S TRSP1. About 5 µg of anti-rabbit HA (Santa Cruz) plus PBS (1x, 0.001% Tween 20) were incubated with Dynabeads ® Protein G [size: 1 ml (30mg/ml) Novex by Life Technologies, Cat# 10003D] for 1 h at 4ºC. 200 µg of total proteins (10x amount used in western blot) were incubated with antibodies and beads with rotation for 2 h at RT. Immuno-complexes were captured on magnet and washed three times with 200 µl PBS 1X. Co-immunoprecipitated proteins were subjected to Western Immunoblot analysis as per regular protocol (anti-Flag, 1:1000). PVDF membrane was stripped and probed with anti-HA antibody (Santa Cruz), 1:1000. The protein bands were visualized by Chemidoc MP imaging system-Biorad.

N. Immunostaining

Hela cells were grown onto 12-well costar culture plates on coverslips at sub confluence (~60%) level and maintained in Dulbecco's modified Eagle's medium (DMEM) having 10% fetal bovine serum (FBS). Transfections were done using polyethylenimine (PEI-sigma). 5 µg of DNA was diluted in 150 µl of serum free DMEM medium and 6 µl of PEI was added into an Eppendorf in a ratio of 1:3 DNA (WT, p.E154Vfs*99 CSRP1, pCMV3-HA-TRSP1, and pCMV3-HA-p.R311S TRSP1) to PEI. Hela cells were fixed in 4% paraformaldehyde (PFA). Nonspecific binding was blocked with 3% Bovine serum Albumin (BSA) in 0.2% PBS-Tween20 (PBT) and

primary mouse monoclonal antibody (1:250) against flag-tag (OCTA-probe H-5) and primary rabbit monoclonal antibody (1:250) against HA-tag (HA-probe Y-11) was applied overnight at 4°C. Secondary anti-mouse or anti-rabbit Biotinylated species-specific whole antibody (from donkey), (GE healthcare UK limited) respectively, were diluted 1:500 and added for 1 hour at room temperature. Cells were washed three times with PBT and incubated for 1 hour at room temperature with Streptavidin Texas Red full length (abcam) diluted 1:500. Cells were washed three times with PBT and incubated with Hoechst staining for the nucleus, diluted 1:30 in water, for 15 minutes. The cells were then mounted on a rectangular slide containing an anti-fading agent DABCO (Sigma-Aldrich). The slides were examined using the Olympus BH-2 or confocal microscope at the molecular core facility in the faculty of medicine-AUB.

O. Immunohistochemistry

Continuous sections that are 5 μm thick were prepared from each formalin-fixed, paraffin embedded tissue. Immuno-histochemical staining was performed to evaluate the expression of CSRP1. All sections on the slides were dewaxed and rehydrated with xylene and graded alcohol, then dripped 3 % hydrogen peroxide on them to quench endogenous peroxidase. Afterwards, high-temperature antigen retrieval was carried out in citrate buffer (pH 6.0) in a microwave oven to enhance immunoreactivity, followed by 3 % Bovine Serum Albumin (Sigma-Aldrich) in 0.2% PBT to reduce the non-specific bindings. Primary rabbit-antibody against CSRP1 (ab70010, abcam, 1:100) was applied to the sections respectively and incubated overnight at 4 °C under humid conditions. Subsequently, slides were incubated at RT with 1:250 secondary anti-rabbit biotinylated species-specific whole antibody (from donkey), (GE healthcare UK

limited) and 1:250 streptavidin–peroxidase conjugate. The antibody-specific binding was visualized with 3, 3-diaminobenzidine solution (DAB -Sigma Aldrich). Finally, slides were counterstained with Methyl Green and mounted. PBS was used as a negative control by replacement of the relevant primary antibody. Images were taken using Laser Microdissection System Leica LMD6500 microscope at American University of Beirut. The specificity of anti-CSRP1 was detected using Rabbit Ig-G, Polyclonal isotype Control for negative control.

P. Statistical analysis

The significance of luciferase assay was studied using Students' T-test with the significance defined as defined as $p < 0.05$ (*), or $p < 0.01$ (**).

Table 7. List of plasmids used.

pCEP4-Flag-CSRP1	pCMV3-HA-TRSP1; pCMV3-HA-p.311STRPS1
pCEP4-Flag-p.E154Vfs*99 CSRP1	PGL3-luc-VEGF
pCGN-HA-GATA4	PGL2-luc-NOS3
pCGN-HA-GATA5	pXP2-luc-NPPA
pCGn-HA-GATA6	pCEP4-HA-TBX5

Table 8. List of primers used.

Gene	Forward Primer	Reverse Primer
WT CSRP1	TGCTCAGCCGAACTGGGAGGAGGCAAG	CAAGGCCTTTGCCACAAGGCCTTTGCCACACTGGC
p.E154Vfs*99 CSRP1	ACTGGTGGCAGCCATTAACT	CTTTCCCCCTCTGGTCTTGT
mCSRP1 cDNA	ATGAGGAAGCCCCTGGAC	AAGCAGGACTTATGCCAGGA
p.R311S TRSP1	TGGGACCTATGATGTGC	TTTAGTAAAACAGGGCTTGAAGAATTGATG

Q. Antibodies

Rabbit polyclonal antibody to CSRP1 Cat#:ab70010; Goat anti-mouse antibodies (HRP) Cat# ab6789; and goat anti-rabbit antibodies (HRP) Cat# ab97051 were purchased from Abcam. Mouse monoclonal antibody against flag-tag (OCTA-probe H-5) Cat# sc-166355 and rabbit polyclonal antibodies against HA-tag (HA-probe Y-11) Cat# sc-805 were from Santa Cruz. Mouse Biotinylated species-specific whole antibody (from donkey), Cat#: LRPN1001V and rabbit biotinylated species-specific whole antibody (from sheep); Cat# RPN1004V were from GE healthcare UK limited. Streptavidin Red full length Cat#ab136227 was from abcam. Chromeo ^{TM} 488 streptavidin green was from Santa Cruz; Cat# sc-364698. Goat anti-rabbit Alexa fluor 488 was from Invitrogen Cat#A11008. Rabbit Ig-G, Polyclonal isotype Control Cat# abcam 27478 was from abcam.

CHAPTER IV

A NOVEL ROLE FOR *CSRP1* IN A LEBANESE FAMILY WITH BOTH CARDIAC DEFECTS AND POLYDACTYLY: A POTENTIAL DIGENIC CARDIAC EFFECT IN CONJUNCTION WITH *TRPS1*

A. Abstract

Despite an obvious role for consanguinity in congenital heart disease, most studies fail to document a monogenic model of inheritance except for few cases. We hereby describe a Lebanese family from a consanguineous marriage between two first-degree cousins. Out of the 7 conceived children, 2 died *in utero* of unknown causes. Of the remaining 5 children, 3 have congenital heart disease, and 4 have polydactyly.

Targeted exome sequencing identified a heterozygous duplication of a 14 nucleotides fragment in *CSRP1*, causing a frameshift mutation at position 154 of the protein. Genotyping family members showed that this mutation is inherited from the father, and segregates only with the CHD phenotype. The variant was neither found in 200 exomes of Lebanese origin, nor in the exome and genome databases. The *in vitro* characterization of the mutation shows no effect on the cellular localization of the protein; however, it dramatically abrogates its transcriptional activity over cardiac promoters like *NPPA*. In addition, it differentially inhibits the physical association of CSRP1 with SRF, GATA4, and with the newly described partner herein TBX5.

Whole exome sequencing failed to show any potential variant linked to polydactyly, but revealed a novel *TRPS1* missense mutation inherited from the healthy mother, and segregating only with the cardiac phenotype. Both TRPS1 and CSRP1 physically interact, and the mutations in each abrogate their partnership.

Our findings add fundamental knowledge into the molecular basis of congenital heart defects and propose the di-genic model of inheritance as responsible for such malformations.

B. Introduction

Congenital heart defects arise during pregnancy, and are subsequently the most prevalent birth defects worldwide(Hoffman & Kaplan, 2002). They affect chamber and valve formation and function leading to different phenotypes referred to as Congenital Heart Disease(Kang et al., 2010), the major cause of neonatal morbidity and mortality in humans. CHD accounts for one third of all main congenital defects with variable prevalence crosswise countries. In Lebanon, the incidence of infants born with CHD between 1980 and 1995 was 11.5 per 1000 live births(Bitar et al., 1999), and twenty per cent of those patients were found to be from first degree cousin mating(Nabulsi et al., 2003). Although many studies have attempted to establish a relationship between CHD and consanguinity, the significance of this association and its precise nature is still unclear. So far, at least 50 human disease genes have been associated with CHD, however, a small set of developmental genes (for example, *NKX2.5*, *GATA4* and *NOTCH1*) harbor the majority of these CHD-associated mutations(Fahed et al., 2013). By understanding the interaction partners, transcriptional targets, and upstream activators of these core cardiac transcription factors, additional information about normal heart formation and further insight into genes and pathways affected in congenital heart disease would emerge.

Mutations in genes encoding LIM domain proteins have been rarely associated to cardiac morphogenesis or CHD. The LIM domain contains a conserved double zinc

finger motif that is evolutionary conserved and is found in a variety of proteins displaying distinct biological roles(Schmeichel & Beckerle, 1994). The LIM domains have been observed to act as a mediator of protein–protein interactions in the cytoplasm and the nucleus. These interactions with specific protein partners are now known to influence its subcellular localization and activity(Camarata et al., 2006; Kadrmas & Beckerle, 2004; Khurana et al., 2002). Many LIM proteins that were initially identified as cytoskeleton-associated proteins, such as members of the cysteine-rich protein (CRP) families, four-and-a-half LIM (FHL), PINCH and Zyxin are recognized to shuttle between the cytoplasm and nucleus of the cell to influence gene expression (Cattaruzza, Lattrich, and Hecker 2004; Chang et al. 2003). This dual localization is due to the presence of a putative nuclear targeting signal (KKYGPK) that has been identified in the glycine-rich regions of the CRPs. In humans, three CRP-family members (group two of LIM domain proteins) have been identified which are: CRP1, CRP2 and CRP3/MLP(Henderson, Brown, Richardson, Olson, & Beckerle, 2002; Pomiès et al., 1997; R. Weiskirchen et al., 1995). CRPs are small proteins, 22 kDa in size, and contain two functional LIM domains that are linked to glycine-rich repeat. CRP family members play a role in terminal differentiation in vertebrate muscle development. CRP1 and CRP2 are prominent in smooth muscle; and CRP3 is expressed in striated muscle (R. Weiskirchen et al., 1995). In the cytoplasm, all three proteins interact with α-actinin (an actin cross-linking protein) and are associated with the actin cytoskeleton, and were also shown to interact with the adhesion plaque LIM protein domain Zyxin(Sadler et al., 1992). In embryonic development CRP1 participates in the formation of heart, and its downregulation alters cardiac-committed mesodermal cell migration resulting in *cardia bifida* in zebrafish(Miyasaka et al., 2007). CRP1 was shown to interact with Wnt

signalling pathway components Dishevelled and Diversin. CRP1 inhibition leads to irregular cell movement in convergent extension resulting in anomalies in midline structures. In adult mouse, CRP1 is expressed in the smooth muscle cells of cardiac arteries. CRP1 and CRP2 were shown to function through coordinated docking of Serum-Response Factor (SRF) to the N-terminal LIM domain and GATA factors, specifically GATA4 and GATA6, to the C-terminal LIM domain(B Lilly, Olson, & Beckerle, 2001; Brenda Lilly et al., 2010). The strong expression of many smooth muscle-differentiation markers is stimulated by this ternary complex of SRF–CRP–GATA, whereas the pairwise combinations have much less impact on gene expression(Chang et al., 2003b). In the cytoplasm, CRPs that are associated with the actin cytoskeleton might function as sensors to assess the physiological status of the contractile machinery. *Csrp2* inactivation results in obvious changes in cardiac ultrastructure, although mice with non-functional CRP2 were otherwise viable and fertile(Brenda Lilly et al., 2010). CRP3/MLP which was first isolated as a cDNA from rat skeletal muscle is a positive regulator of myogenic differentiation. The overexpression of CRP3 in myoblasts enhances skeletal myogenesis while its silencing prevents them from exiting cell cycle and block terminal differentiation(Arber et al., 1994).

We have recently identified a large Lebanese family with CHD and polydactyly composed of the consanguineous marriage between two first-degree cousins. Out of the 7 conceived children, 2 died *in utero* at the ages of 6 and 9 months of unknown causes. Of the remaining 5 children, 3 have congenital heart disease (ventricular septal defect, atrial septal defect, and patent ductus arteriosus), and 4 have polydactyly (2 have both). We performed targeted exome sequencing of 119 candidate genes (Supplementary

Table 1) and identified in all affected probands a novel heterozygous frameshift mutation in *CSRP1* inherited from the unaffected father, and a damaging missense mutation in *TRPS1* inherited from the unaffected mother. We hypothesized that a digenic mode of inheritance explains the occurrence of CHD in this family.

C. Materials and methods

1. Patient recruitment and clinical examination

The study was approved by the institutional review board at the American University of Beirut (protocol number: Bioch.GN.01). All patients and family members signed an informed consent form before being enrolled in the study. Genetic analyses and return of genetic data were performed in accordance with protocols approved by the Partners Human Research Committee. A total of 20 individuals from the same family were enrolled. Standard clinical evaluation included a comprehensive physical exam, electrocardiography (ECG), and two-dimensional (2D) transthoracic echocardiography (TTE) with color Doppler.

2. Genetic analysis

Peripheral venous blood was collected from all family members, and DNA extraction was performed using the Qiagen Blood-Midi kit (Qiagen Science Inc., Germantown, MD, USA), following the manufacturer's protocol. DNA quantification was performed using the NanoDrop (Thermo Fisher Scientific Inc., Waltham, MA, USA) at the molecular core facility at AUB. Targeted DNA sequencing was done at Harvard as previously described. One microgram of coded DNA samples from both parents, the proband, four of her siblings, and three of her cousins were shipped to

Macrogen (South Korea) where exome sequencing was performed using the Agilent V6 Sureselect target enrichment capture system on a HiSeq2000 platform from Illumina (San Diego, USA). Primary analysis was done at Macrogen. Generated Fastq files were mapped to the reference genome using the Burros-Wheeler Alignment Tool (BWA), and the Genome Analysis Toolkit (GATK) was used for variants call, while the SnEff software was used to annotate the variants. Sanger sequencing was used to confirm the genotype the *CSRP1* and *TRPS1* variants in all available family members. Briefly, amplification by polymerase chain reaction (PCR) was done using the Phusion polymerase high-fidelity master mix (F-548S) on a Pico machine (Finnzymes, Espo, Finland), and the amplicons were resolved on a 1.5% agarose gel. Gel purification was performed using the Gel Extraction kit following the manufacturer's protocol (peqGOLD Gel Extraction Kit, PeqLab, Erlangen, Germany). The purified bands were quantified using a NanonDrop (Thermo Fisher Scientific Inc., Waltham, MA) and examined by gel electrophoresis to ensure quality. DNA sequencing was carried out on an ABI 3500 machine at the molecular core facility at the American University of Beirut, followed by analysis using the data collection software from Applied Biosystems Inc. (Foster City, CA).

3. Cell lines and plasmids

HEK293 cells (Human Embryonic Kidney cells), Hela cells (Human cervical cancer cells) were cultured and maintained in Dulbecco’s Modified Eagle Media (DMEM-Sigma, Cat#D0819) supplemented with 10% Fetal Bovine Serum (FBS-Sigma, Cat#F9665), 1% Penicillin/ Streptomycin (Biowest-Cat#L0022-100) and 1%

Sodium pyruvate (Sigma-Cat#S8636). Incubation of cells was carried out in a 5% CO2 humid atmosphere at 37ºC.

Luciferase reporter plasmids PGL3-VEGF-luc, PGL2-NOS3-luc, and PXP2-NPPA-luc were constructed by ligation of PCR–amplified fragments from mouse VEGF, human NOS3 and rat NPPA promoters into eukaryotic luciferase expression vectors PGL3, PGL2 and pXP2 respectively. pCGN-HA-SRF, HA-tagged pCGN-GATA (4, 5, and 6) and pCGN-TBX5 were cloned into eukaryotic expression vector pCGN. pCMV3-HA-TRSP1 was from Sino Biological Inc. (Cat# HG15989-NY). Flag-tagged wild type human CSRP1 was generated by subcloning CSRP1 fragments into expression vector pCEP4 (Invitrogen). Flag-tagged p.E154Vfs*99 CSRP1 mutant was constructed through site-directed mutagenesis in which PCR amplified fragments harboring the p.E154Vfs*99 mutation were ligated into eukaryotic expression vector pCEP4. Site directed mutagenesis was also carried out to introduce p.R311S *TRPS1* mutation into the wild-type (WT) TRPS1. After the ligation of the resulting amplicons, transformation into XL-1 Blue competent bacteria was performed. Finally, the yielded plasmids were extracted and sequenced thereafter in order to confirm the incorporation of the mutations.

4. Antibodies

Rabbit polyclonal antibody to CSRP1 Cat#:ab70010; Goat anti-mouse antibodies (HRP) Cat# ab6789; and goat anti-rabbit antibodies (HRP) Cat# ab97051 were purchased from Abcam. Mouse monoclonal antibody against flag-tag (OCTA-probe H-5) Cat# sc-166355 and rabbit polyclonal antibodies against HA-tag (HA-probe Y-11) Cat# sc-805 were from Santa Cruz. Mouse Biotinylated species-specific whole

antibody (from donkey), Cat#: LRPN1001V and rabbit biotinylated species-specific whole antibody (from sheep); Cat# RPN1004V were from GE healthcare UK limited. Streptavidin Red full length Cat#ab136227 was from abcam. Goat anti-rabbit Alexa fluor 488 was from Invitrogen Cat#A11008.

5. Transfection assays to assess CSRP1 target gene promoters

HEK293 cells were grown and maintained at subconfluence ~60% level in Dulbecco's modified Eagle's Medium (DMEM-Sigma) containing 10% fetal bovine serum. Transient transfections were performed with Polyethylenimine (Sigma). A series of Luciferase assays were performed by transient transfection in combination with pCEP4-Flag-WTCSRP1, pCEP4-Flag-p. E154Vfs*99 CSRP1, PCGN-HA-SRF, pCGN-HA-GATA4, pCGN-HA-GATA4, luciferase reporters (3.5 µg) and empty expression vectors, PCGN, to a balanced total of 1 µg of plasmids per 2 wells of the 12-well plate. The results were normalized to total protein concentration in each well, and were expressed as fold activation. Cotransfection experiments were performed in duplicates and repeated at least three times. Luciferase activity was normalized to baseline reporter gene activity as fold activation, with error bars representing SEM.

6. Protein overexpression and western blotting

HEK293 cells were transiently transfected with epitope-tagged vectors pCGN-HA-GATA4, pCEP4-Flag-WTCSRP1, pCEP4-Flag-p.E154Vfs*99 CSRP1, pCEP4-HA-TBX5, pCGN-HA-GATA5, pCGN-HA-GATA6, PCGN-HA-SRF, pCMV3-HA-WT TRSP1, and pCMV3-HA-p.R311S TRPS1 using Polyethylenimine (Sigma). HEK293 cells were plated in 100 mm corning culture plates until sub confluence ~80%

determined by green fluorescent protein (GFP) transfection assay. After 24 hr, 20 µg of DNA and 35 µl PEI (transfection reagent) were added to an Eppendorf tube holding a total volume of 1 ml DMEM medium. The mixture was vortexed for 10 sec, incubated 20 min at RT, and then applied over the cells. Culture medium was changed after 3 h of transfection. Nuclear extracts from transfected HEK293 cells were obtained as previously described. For immunoblotting, 10 µg of nuclear extracted proteins were mixed with 5X Laemmli Buffer. The protein samples were boiled for 5 min and run on denaturing SDS-PAGE for about 1.5 h then transferred to a PVDF membrane (Amersham, UK) Cat#10600023. The membrane was blocked in 5%TBT (TBS-0.02% Tween 20) skimmed dry milk for 45 min at RT. The membrane was incubated with primary antibodies, anti-Flag or anti-HA (1:1000) overnight at 4°C. On the second day, the membrane was washed three times with TBT and incubated with secondary anti-mouse or anti-rabbit-HRP (1:50000) for 1 hr at RT. Development was done using ECL ™ Western Blotting Detection Reagents (Amersham, GE healthcare, Cat# RPN2106). The protein bands were visualized by Chemidoc MP imaging system-Biorad and quantified using Image J software.

7. Coimmunoprecipitation

After detecting WT CSRP1, p.E154Vfs*99 CSRP1, GATA-4,-5, and -6, SRF, TBX5 , WT TRPS1, and p.R311S TRPS1 proteins by western blot, co-immunoprecipitation assay was done to assess the physical interaction between WT/ p.E154Vfs*99 CSRP1 (Flag-tagged), GATA-4,-5, and -6 (HA-tagged), SRF (HA-tagged), TBX5 (HA-tagged), and WT/p.R2311S TRPS1(HA-tagged). About 5 µg of anti-rabbit HA (Santa Cruz) plus PBS (1x, 0.001% Tween 20) were incubated with

Dynabeads ® Protein G [size: 1 ml (30mg/ml) Novex by Life Technologies, Cat# 10003D] for 1 h at 4ºC. 200 µg of total proteins (ten times the amount used in western blot) were incubated with antibodies and beads for 2 hr at RT. Immunocomplexes were captured on magnet and washed three times with PBS 1X. Coimmunoprecipitated proteins were subjected to Western Immunoblot analysis as per regular protocol (anti-Flag, 1:1000). PVDF membrane was stripped and probed with anti-HA antibody (Santa Cruz), 1:1000. The protein bands were visualized by autoradiography.

8. Immunostaining

Hela cells were grown onto 12-well costar culture plates on coverslips at sub confluence (~60%) level and maintained in Dulbecco's modified Eagle's medium (DMEM) having 10% fetal bovine serum (FBS). Transfections were done using polyethylenimine (PEI-Sigma). 5 µg of DNA was diluted in 150 µl of serum free DMEM medium and 6 µl of PEI was added into an Eppendorf in a ratio of 1:3 DNA to PEI. Hela cells were fixed in 4% paraformaldehyde(Hummel, Li, Pfaffinger, Neven, & Scanu, 1990). Nonspecific binding was blocked with 3% Bovine serum Albumin (BSA) in 0.2% PBS-Tween20 (PBT) and primary mouse monoclonal antibody (1:250) against flag-tag (OCTA-probe H-5) or rabbit polyclonal IgG (HA-probe Y-11) were applied overnight at 4°C. Secondary anti-mouse Biotinylated species-specific whole antibody (from donkey), GE healthcare UK limited, or secondary anti-Rabbit biotinylated species-specific whole antibody (from sheep); diluted 1:500 was added for 1 hour at room temperature. Cells were washed three times with PBT and incubated for 1 hour at room temperature with Streptavidin Texas Red full length (abcam, Cat#ab136227) diluted 1:500. Cells were washed three times with PBT and incubated with Hoechst

staining for the nucleus, diluted 1:30 in water, for 15 minutes. The cells were then mounted on a rectangular slide containing an anti-fading agent DABCO (Sigma-Aldrich). The slides were examined using the Olympus BH-2 microscope at the molecular core facility in the faculty of medicine-AUB.

9. Immunohistochemistry

Continuous sections which are 5 μm thick were prepared from each formalin-fixed, paraffin embedded tissue. Immuno-histochemical staining was performed to evaluate the expression of CSRP1. All sections on the slides were dewaxed and rehydrated with xylene and graded alcohol, then dripped 3 % hydrogen peroxide on them to quench endogenous peroxidase. Afterwards, high-temperature antigen retrieval was carried out in citrate buffer (pH 6.0) in a microwave oven to enhance immunoreactivity, followed by 3 % Bovine Serum Albumin (Amresco Life science, Cat#0332-100G) in 0.2% PBT to reduce the non-specific bindings. Primary rabbit-antibody against CSRP1 (ab70010, abcam, 1:100) were applied to the sections respectively and incubated overnight at 4 °C. Subsequently, slides were incubated with 1:250 secondary anti-rabbit Biotinylated species-specific whole antibody (from donkey), (GE healthcare UK limited, Cat# RPN1004V) and 1:250 streptavidin–peroxidase conjugate, and antibody-specific binding was visualized with 3, 3-diaminobenzidine solution (DAB –Sigma, Cat#D3939-1SET). Lastly, slides were counterstained with Methyl Green and mounted. PBS was used as a negative control by replacement of the relevant primary antibody.

10. Statistical analysis

The significance of luciferase assay was studied using Students' T-test with the significance defined as $p < 0.05$ (*), or $p< 0.01$ (**).

D. Results

A CHD consanguineous multiplex family with congenital heart disease and polydactyly was recruited as part of the Congenital Heart Disease Genetics Program at the American University of Beirut (figure 16). The indexed-patient (III-8, figure 16) presented to the AUB-MC Children's Heart center shortly after birth with a patent ductus arteriosus (PDA) and bilateral postaxial polydactyly on both hands and feet: the PDA was closed using an Amplatzer^{TM} device. The girl passed away at 7 years of age following a severe lung infection. Her brother (III-10) presented with a small perimembraneous ventricular septal defect (VSD), lower extremity bilateral postaxial polydactyly, right hand syndactyly with postaxial finger formation, and left hand postaxial polydactyly with syndactyly in the 4^{th} and 5^{th} digits. Further examination of the core family (Figure 17) showed that the father (II-8) has bilateral postaxial polydactyly on feet and hands while the mother (II-9) is phenotypically normal (Figure 17). Both parents and the unaffected child (III-9) were confirmed upon echocardiography to be CHD free. The parents are first degree cousins, and had recently monozygotic twin-girls; one of them with a patent foramen ovale (PFO) and left-hand postaxial polydactyly (III-11), while the other (III-12) has only an atrial septal defect (ASD) (Figure 17).

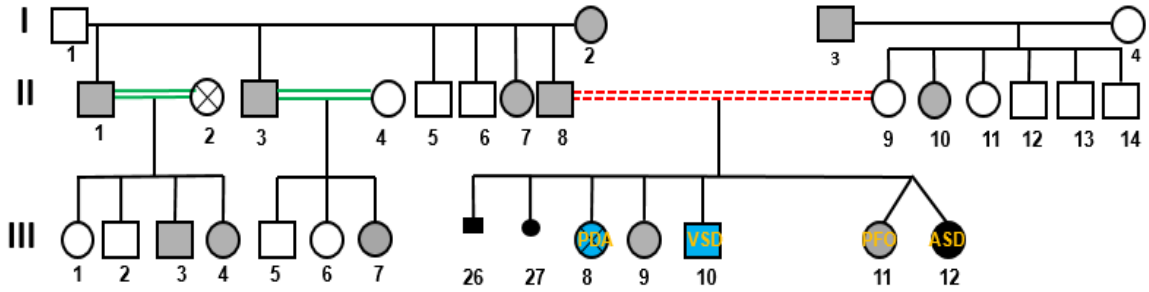


Figure 17. Congenital heart and limb deformities in a large consanguineous Lebanese family. The pedigree shows a three generations' family (roman numbers) with inherited polydactyly (grey symbols) or congenital heart defects (black symbols) amongst its members (arabic numbers). Red and green dotted lines indicate first and second degree cousin marriages respectively. Stillbirths are represented with small dark circles and squares for females and males respectively. Deceased individuals were crossed. VSD: ventricular septal defect, ASD: atrial septal defect, and PDA: patent ductus arteriosus.

1. CHD targeted sequencing: The CSRP1 variant

Targeted sequencing of 119 genes implicated in CHD for probands III-8 and III-10 showed that none of these genes harbor shared rare (MAF<1%) damaging variants except for the *CSRP1* gene.

The variant [NM_004078.2:c.447_460dupTGGCAAAGGCCTTG] is an insertion of a segment of 14 nucleotides chr1:20145445-T>TCAAGGCCTTTGCCA that leads to a frameshift mutation with an extended C-terminal domain of the protein p.E154Vfs*99 (figure 18A). The variant was shared by both probands and was not present in the Genome Aggregation Database (gnomAD) which includes 123,136 exome sequences and 15,496 whole-genome sequences from unrelated individuals, nor in 200 Lebanese patients with CHD screened using the same approach. The variant was confirmed by Sanger sequencing (figure 18B).

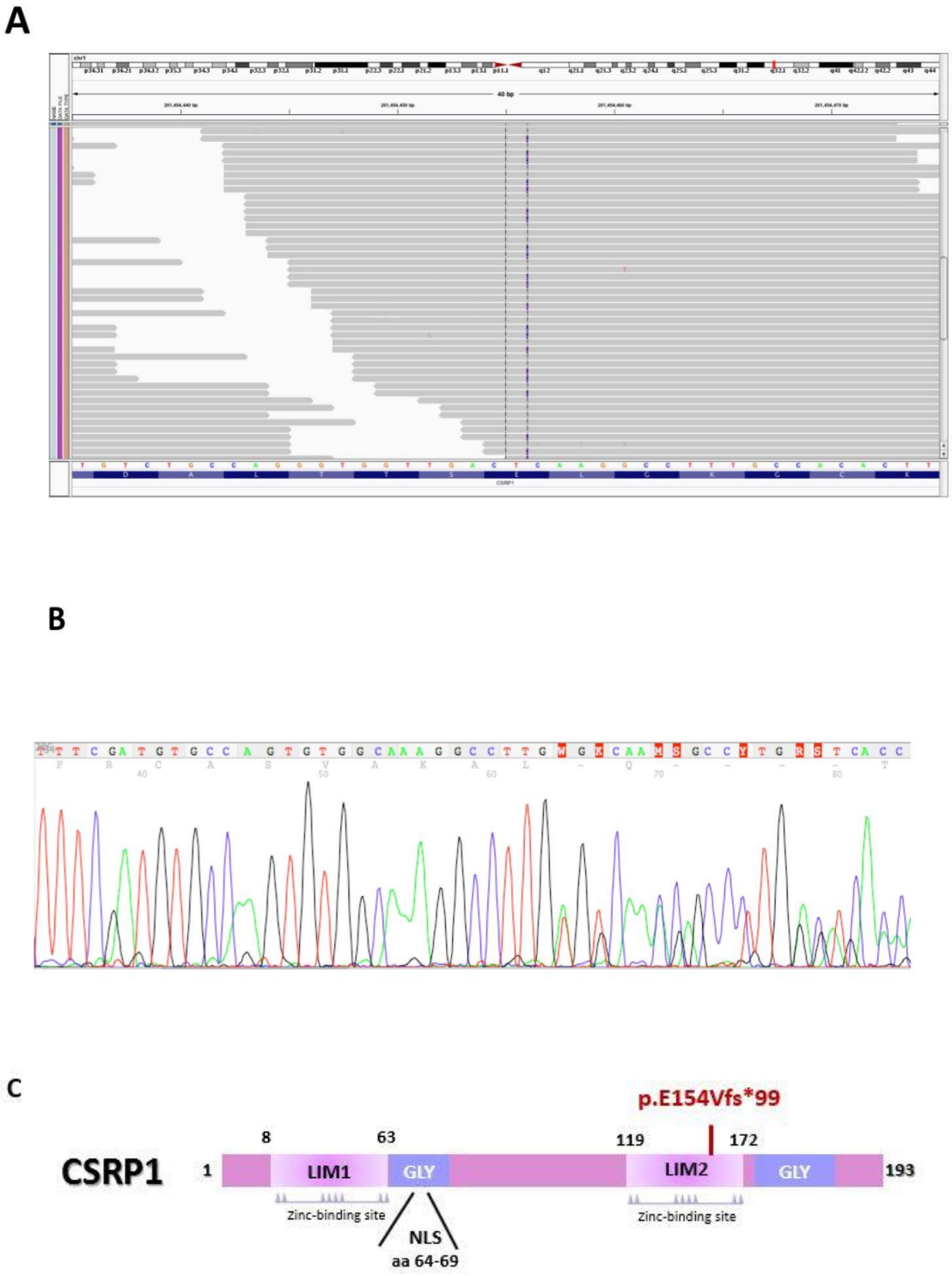


Figure 18. Sequencing results of the affected patients along with CSRP1 schematic representation. A. Integrative genomics viewer (IGV) visualization of the targeted exome sequencing shows an insertion (Blue line) in Csrp1 gene. **B.** Sanger Sequencing of the *CSRP1* gene confirmed the 14 nucleotides duplication (TGTGGCAAAGGCCT) in exon 5. **C.** Schematic representation of the mutation that abrogates the second LIM domain of the protein.

Genotype-phenotype analysis across the extended pedigree did not support any role for CSRP1 in the polydactyly phenotype, but suggested a potential role in the cardiac phenotype observed in the core family, since the third affected child with ASD III-12 carried the variant, whereas proband III-9 who has no cardiac phenotypes did not (Table 9). The variant is inherited from the father who has a normal heart on echocardiography. This suggested that this variant could be involved in the cardiac phenotype observed in the family.

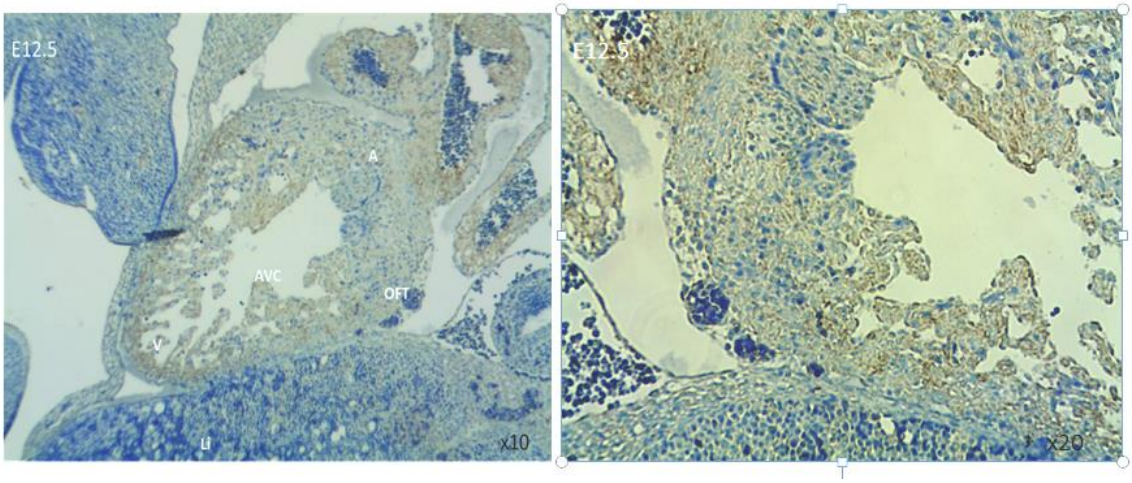
Table 9. Genotype of Family Members for CSRP1 and TRPS1 variants. (-/- : no mutation; +/-: heterozygote mutation; N/A: non-applicable).

Family Member/ Mutation	CSRP1(p.E154Vfs*99)	TRSP1(p.R311S)
I-1,2,3,4	NA	NA
II-1	-/-	NA
II-2,3,4,5,6,7	NA	NA
II-8	+/-	-/-
II-9	-/-	+/-
II-10	-/-	NA
II-11,12,13,14	NA	NA
III-1,2	NA	-/-
III-3,4	-/-	NA
III-5,6,7	NA	NA
III-8	+/-	+/-
III-9	-/-	-/-
III-10	+/-	+/-
III-11	+/-	+/-
III-12	+/-	+/-

We retrieved the coding sequence of the human *CSRP1*. Our analysis indicates a duplication of 14 nucleotides in exon 5 (Figure 18A) leading to a reading frameshift that disrupts the 2nd LIM domain of the protein (NP_001180500) and extending its C-terminus (figure 18C). The mutated protein harbors 253 amino acids instead of 193 suggesting a potential conformational, structural, and functional change.

2. Cardiac expression and cellular localization of CSRP1

Before characterizing the impact of the mutation on the protein cellular localization and transcriptional activity, we did look at the expression of the protein during heart development in mice to correlate it with the phenotype(s) observed in the affected individuals. Results of immuno-staining showed a strong expression of the protein in the heart at all stages of development starting as early as E12.5 and onwards with a strong nuclear expression in all cardiac compartments, but not in the valves (figure 19). Of note the absence of the expression of the protein in the endocardial cells and the progressive increase in the cytoplasmic localization of the protein whereby the newborn heart mice have mainly cytoplasmic CSRP1.



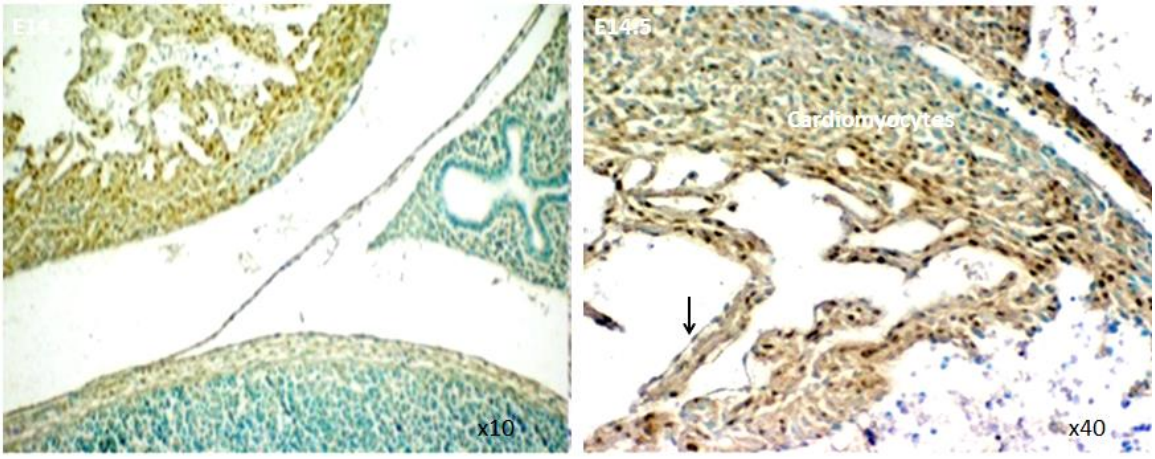


Figure 19. Frontal sections of mouse embryos at different developmental stages (E12.5 and E14.5). CSRP1 shows a higher expression in nuclei (dark brown dots) of cardiomyocytes compared to the cytoplasm (light brown color) at embryonic stages. Nuclei were counter-stained with methyl green. Photos were taken at magnification 10x, 20x and 40x respectively. (A: aorta; OFT: outflow tract; v: ventricle; AVC: atrioventricular canal).

3. Characterization of the transcriptional activity of the CSRP1 mutated protein

In order to assess the impact of the p.E154Vfs*99 mutation on structural and functional properties of CSRP1 protein, site-directed mutagenesis was carried out on the coding region of the human CSRP1 cDNA subcloned into a Flag-tagged plasmid. Both plasmids were sequenced before transiently expressing them into Hela and HEK293 cells. Both the wild type and mutated proteins were equally produced as assessed by western blot analysis of nuclear proteins extracted from these cells (figure 20A). Cellular localization was assessed in Hela cells which showed that both the wild type and mutated proteins are present in both the nuclei and cytoplasm of the transfected cells with no substantial differences (figure 20B).

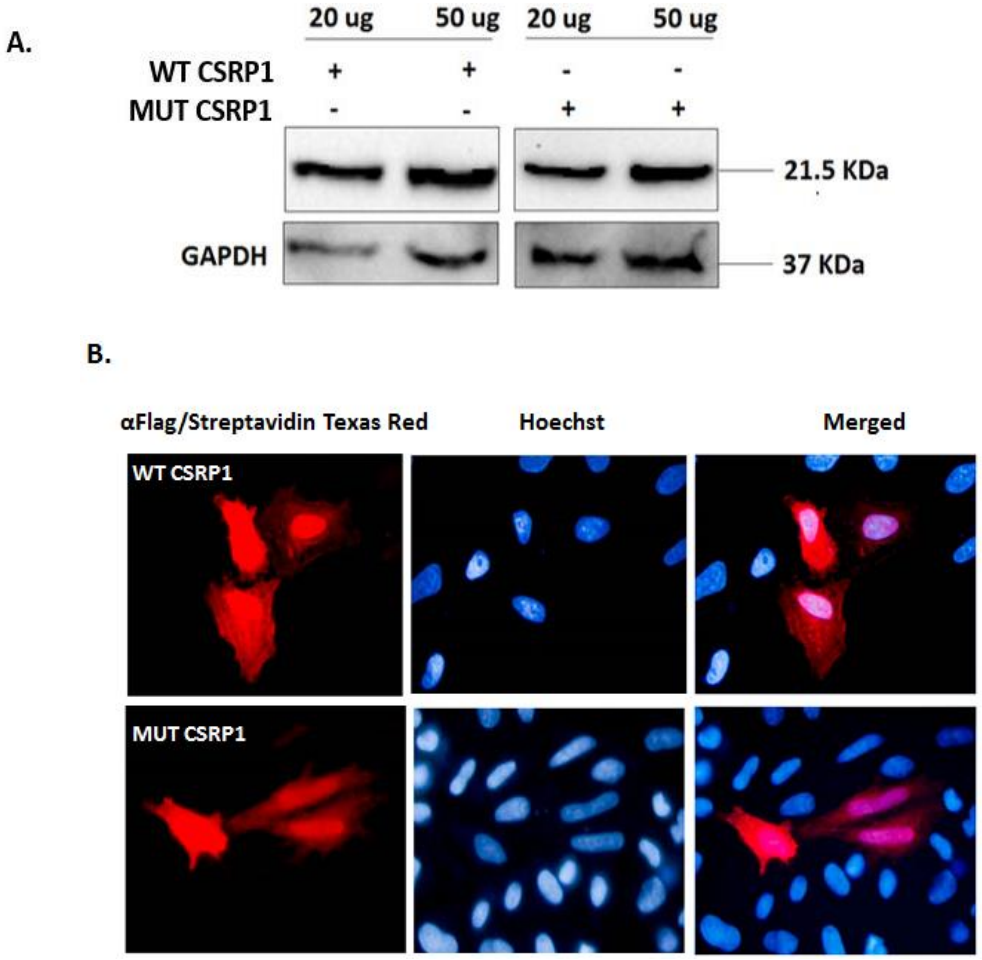


Figure 20. Cellular localization of WT and MUT CSRP1 proteins. (**A**) Nuclear extracts from transfected HEK293 cells with either WT CSRP1 or MUT CSRP1 were resolved on an SDS-PAGE. Immunoblotting using anti-Flag antibody showed equal amounts of expressed proteins at 20 µg and 50 µg. Anti-GAPDH was used as a loading control. (**B**) Immunofluorescence of Hela cells transfected with 20 µg of WT CSRP1 and MUT CSRP1 plasmids. The localization of CSRP1 was visualized using anti-Flag antibody followed by biotinylated anti-mouse antibody and then fluorescent Streptavidin Texas Red. Nuclei were stained blue with the Hoechst 33342 dye. CSRP1 (WT or MUT) showed cytoplasmic and nuclear localization (red color).

Although CSRP1 is predicted to act as a transcription factor, no previous studies have shown a direct activity of this protein on promoter regions of genes. In order to assess the effect of p.E154Vfs*99 mutation on the function of CSRP1, HEK293 cells were transiently transfected with increasing concentrations of the plasmid encoding either the wild type CSRP1 or the variant along with a fixed dose of either one of the

following cardiac-enriched promoters fused to Luciferase: NPPA, VEGF and NOS3. While the wild type protein was able to activate the promoters in a dose-dependent manner reaching up to 6 fold over the NOS3 promoter, the mutation completely abolished this activity (figure 21). NPPA is an early biomarker of cardiac development and it is the most studied example in literature for gene regulation. In addition, NPPA acts as a platform to study regulatory processes during development. It is highly expressed in differentiating chambers during cardiac formation(Houweling et al., 2005). Although NPPA knockout mice are viable, yet they suffer of high blood pressure and cardiac hypertrophy in response to volume overload(Mori et al., 2004). NOS3 is also essential for cardiac development. NOS3 knockout mice suffered of atrial septal defects(Feng et al., 2002). In addition, it was shown that an endocardial pathway requires NOS3, TBX5, and GATA4 for atrial septal formation(Nadeau et al., 2010). As for VEGF, it is also essential for normal cardiac development(Richards & Garg, 2010). The deletion of one VEGF allele in mice leads to embryonic lethality(Carmeliet et al., 1996; Ferrara et al., 1996). Moreover, VEGF upregulation in hypoxia inhibits endocardial to mesenchymal transformation which is essential for valve formation(Dor et al. , 2001). It is essential to mention that NOS3 and VEGF are both expressed in vasculature(Coultas, Chawengsaksophak, & Rossant, 2005; Förstermann & Münzel, 2006) as CSRP1.

The weak transactivation properties of the CSRP1 protein coupled to the lack of a bona fide binding site for this class of LIM proteins prompted us to look at the effect of the mutation on its interacting partners.

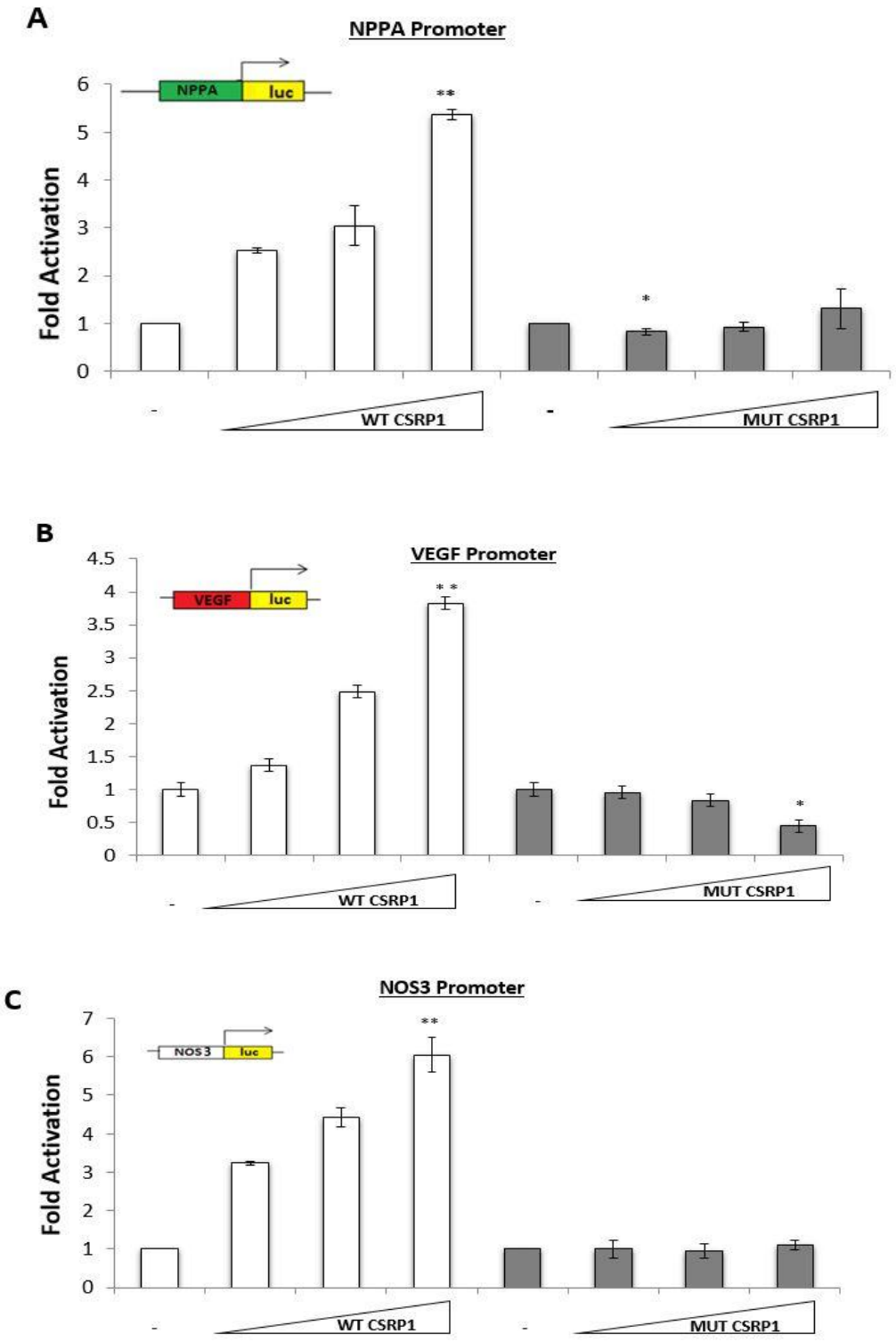
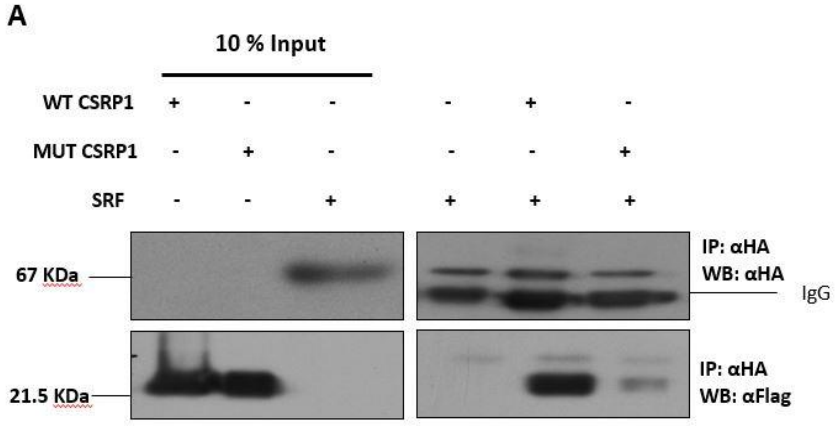


Figure 21. Transcriptional activity of WT and/or MUT CSRP1. WT or MUT CSRP1 were transiently transfected along with 3.5 µg of NPPA, VEGF, and NOS3-luciferase promoter in HEK293 cells. Relative luciferase activities were presented as fold changes. The data represent the means of 3 independent experiments done in duplicates and the values are ± SE. P-value was assessed used Students' T-test. Significance $p<0.01$ is indicated by an (**) while $p<0.05$ is indicated by (*); significance is tested relative to control. WT CSRP1 significantly activates the promoters while MUT CSRP1 does not show any significant activation. WT CSRP1, but not MUT activates the promoters in a dose response manner. The triangle represents an increasing dose of the WT and MUT CSRP1 (200 ng, 400 ng and 600 ng respectively).

*4. The p.E154Vfs*99 mutation alters the physical and functional interaction between CSRP1 and SRF*

CRSP1 was previously shown to be recruited preferentially by the serum response factor (SRF) protein to promoter regions of target genes involved in smooth muscle cells differentiation(Chang et al., 2003b). We thus assessed the effect of the mutation on this interaction by co-immunoprecipitation assays on HEK293 cells transiently overexpressing both proteins. The results showed that the mutation drastically inhibited the CSRP1/SRF interaction by 82% (figure 22A). The functional interaction was subsequently tested in co-transfection assays which show that both CSRP1 and SRF can synergistically activate the NPPA, VEGF, and NOS3 promoters up to 100, 22.3, and 14 folds respectively (figure 22B-D), and that the mutation completely inhibit this synergy over the VEGF, and NOS3 promoters while drastically inhibiting it by 50% over the NPPA promoter.



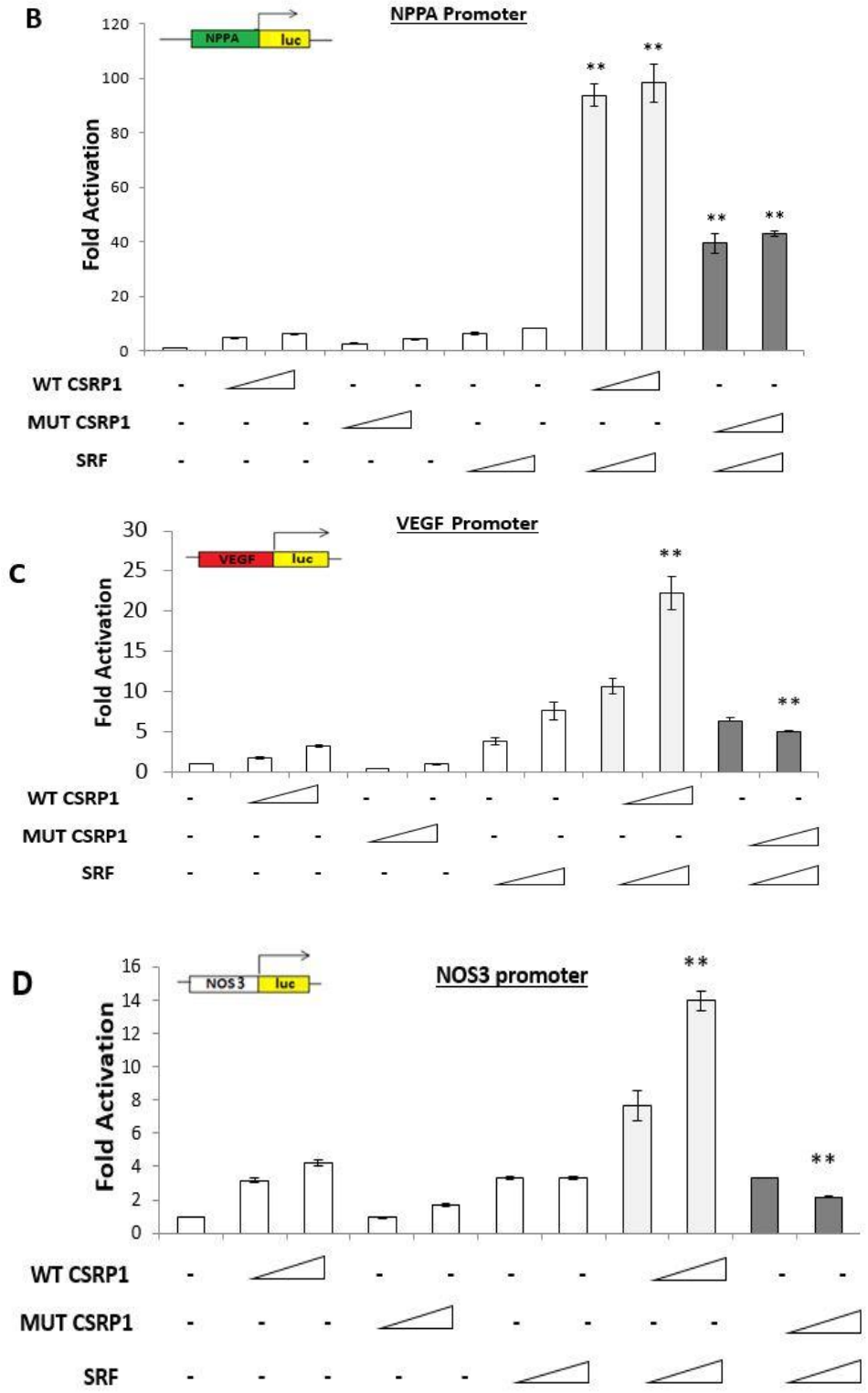
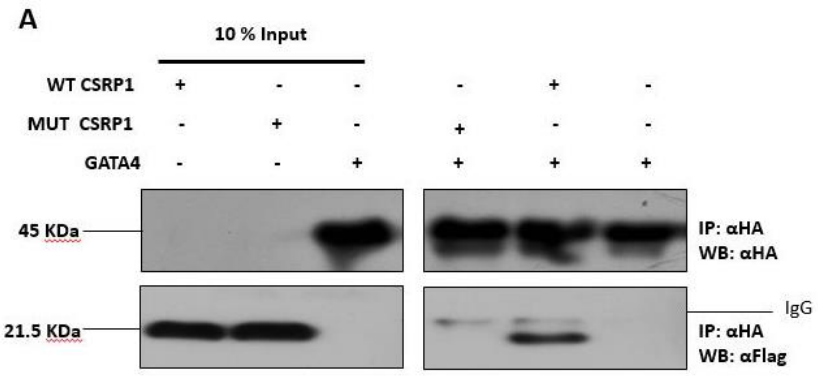


Figure 22. The p.E154Vfs*99 mutation abolishes the physical and functional interaction between CSRP1 and SRF. (A) The amount of the proteins used for immunoprecipitation was ten times that used in western blot. Nuclear lysates of CSRP1 protein were immune-precipitated with anti-HA antibody and CSRP1 protein was detected using anti-Flag antibody (arrows). After membrane stripping, subsequent western blot analysis using anti-HA was performed to detect SRF protein (arrowheads). (B, C, D) WT or MUT CSRP1 were transiently co-transfected with SRF along with 3.5 µg of NPPA, VEGF, and NOS3-luciferase promoters respectively in HEK293 cells. Media was changed 3 h post transfection and cells were harvested for luciferase assay after 36 h. Relative luciferase activities were presented as fold changes. The data represent the means of 3 independent experiments done in duplicates and

the values are ± SE. P-value was assessed used Students' T-test and significance $p<0.01$ is indicated by (**), while $p<0.05$ is indicated by an (*) ; significance of synergy for WT is tested relative to the sum of individual activations, while that of mutant is tested relative to synergy . The triangle represents an increasing dose of the WT and MUT CSRP1 (400 ng and 600 ng respectively) and SRF (200 and 400 ng).

*5. CSRP1/GATA4 interaction altered by the p.E154Vfs*99 CSRP1 variant*

GATA transcription factors and LIM-domain proteins (including CSRP1) have comparable zinc finger motifs through which they heterodimerize(Chang et al., 2003b). Thus, we assessed the effect of the mutation by co-immunoprecipitation assay on HEK293 cells transiently expressing WT CSRP1, p.E154Vfs*99 variant, and GATA4, or GATA5, or GATA6. Our results show that only GATA4 and GATA6 readily interacts with CSRP1 while GATA5 does not (figure 23). More importantly the CSRP1 variant dramatically inhibited the GATA4/CSRP1 interaction by (up to 85%) but has a mild effect on the CSRP1/GATA6 interaction, increased by 15% (figures 23A and 23B). Concurrently, the functional interaction between GATA4 and CSRP1 was also tested in co-transfection assays. The synergy between the two proteins was completely lost on the cardiac enriched promoters NPPA, VEGF, and NOS3 promoters going down from 22 folds to 5 (figure 24).



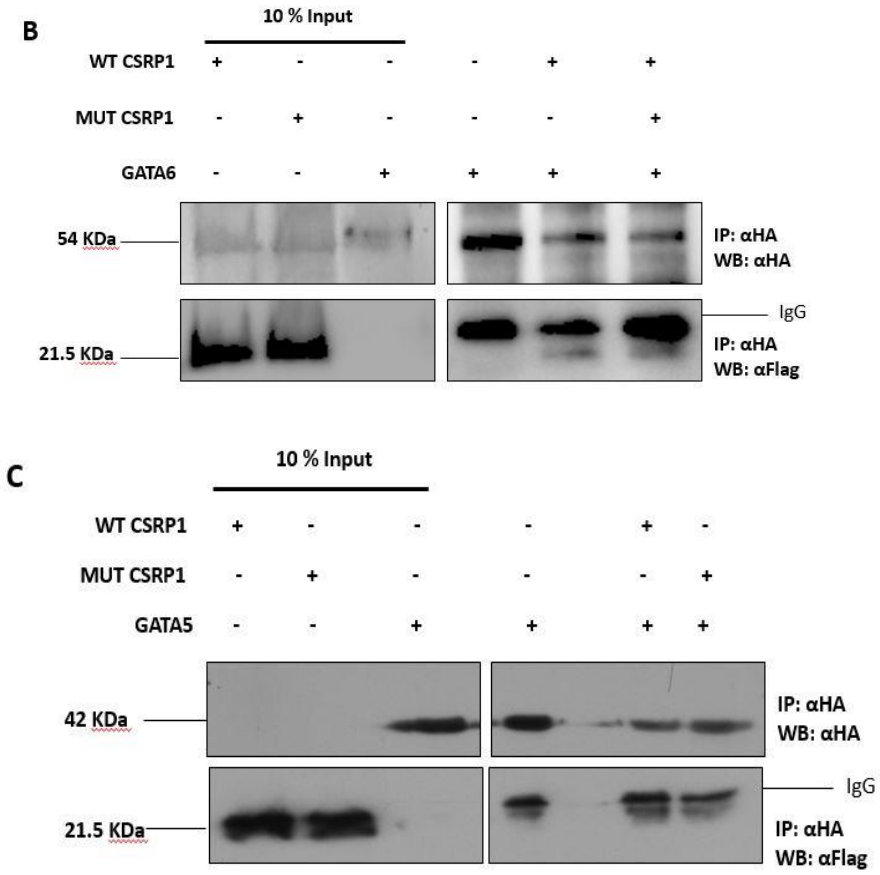
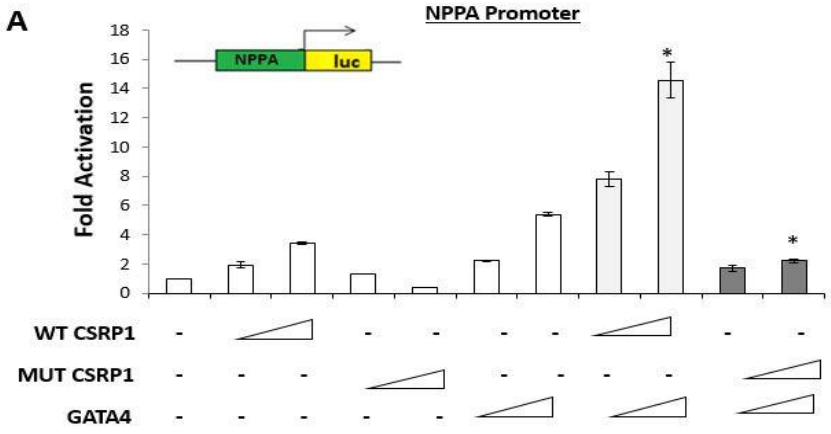


Figure 23. Selective inhibition of the GATA4 physical interaction with CSRP1 by the p.E154Vfs*99 mutation. (A, B, C) The amount of the proteins used for immunoprecipitation was ten times that used in western blot. Nuclear lysates of CSRP1 protein were immunoprecipitated with anti-HA antibody and CSRP1 proteins were detected using anti-Flag antibody. After membrane stripping, subsequent western blot analysis using anti-HA was performed to detect GATA4, 5, and 6 proteins.



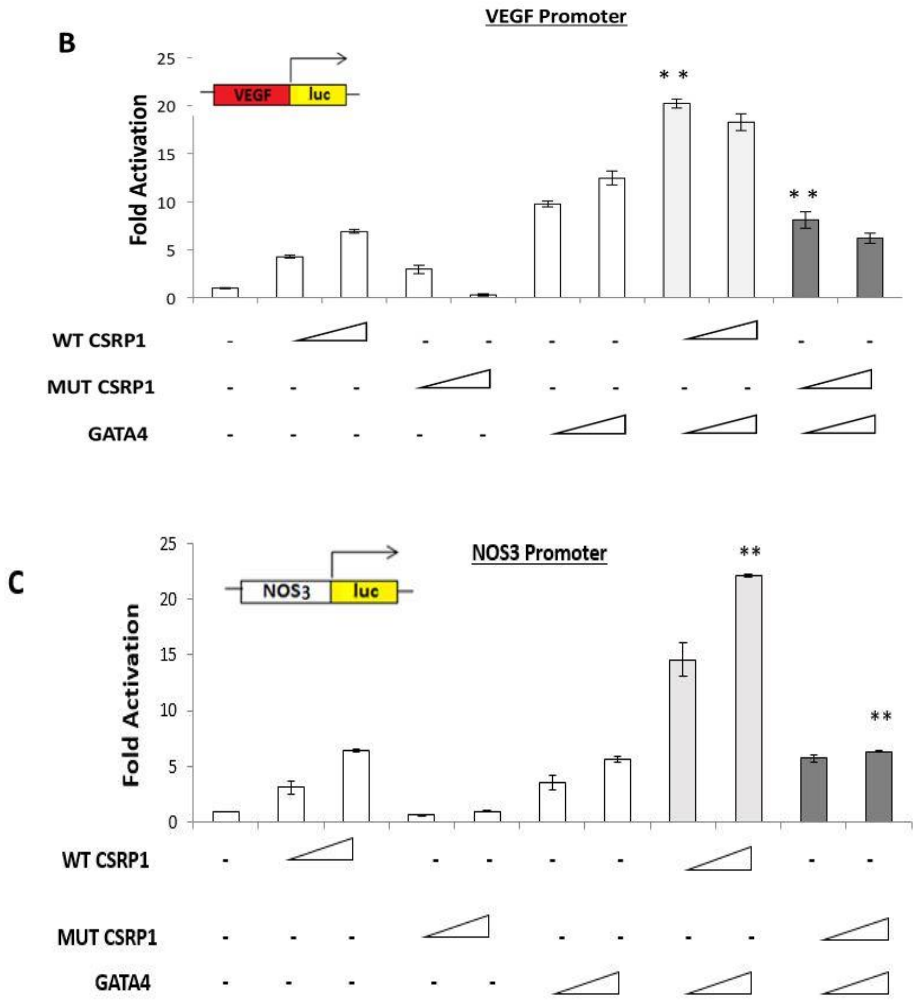
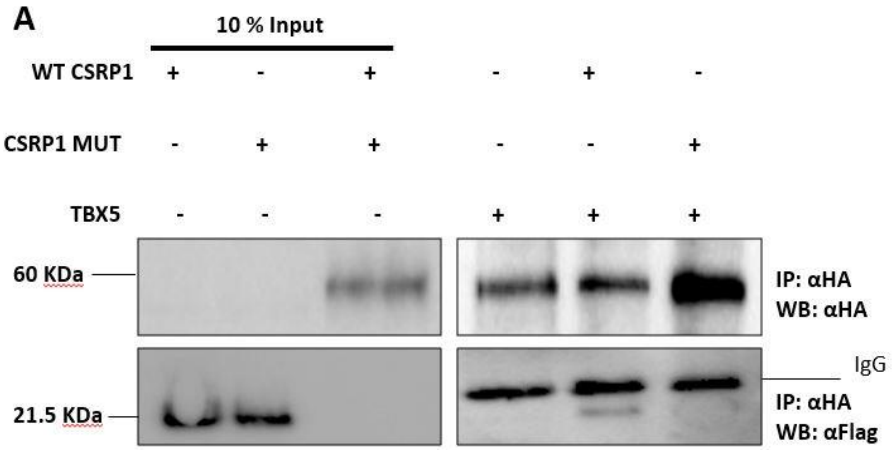
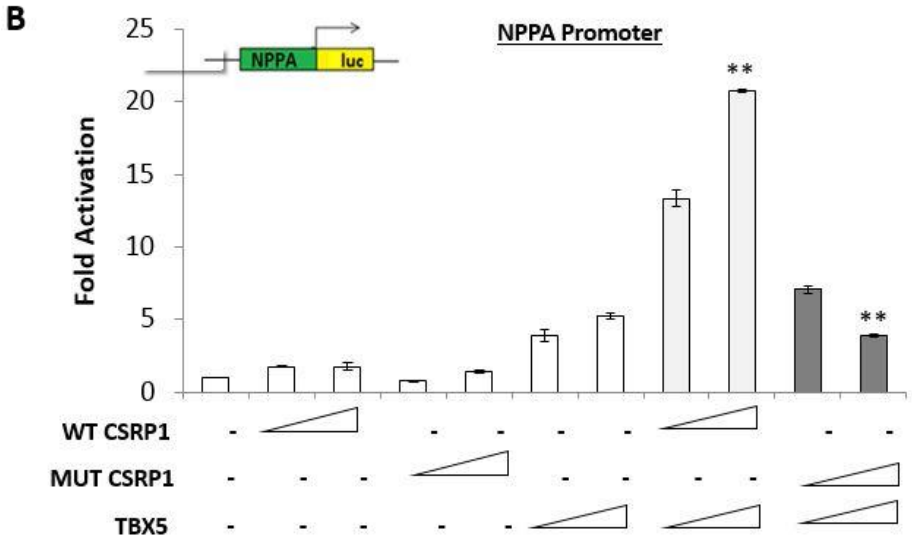


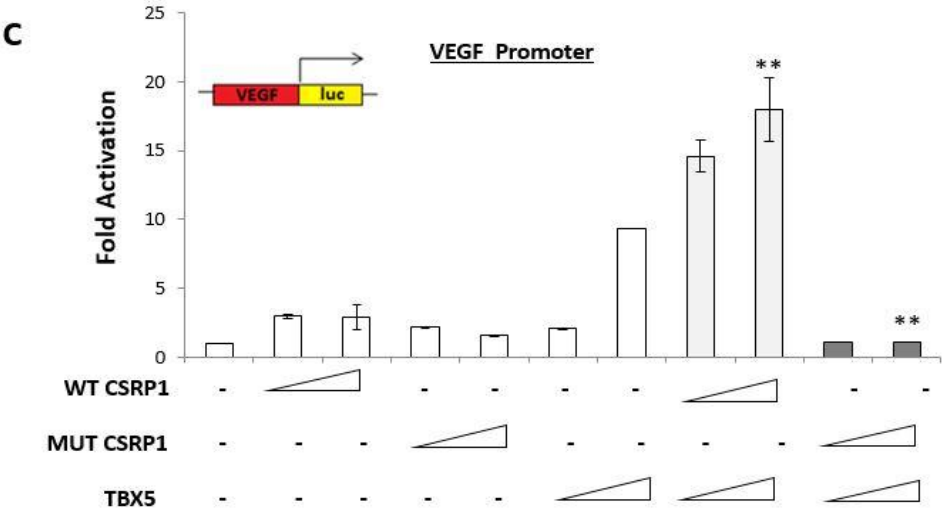
Figure 24.Transcriptional activity of WT and/or MUT CSRP1 along with GATA4. (A, B, C) WT or MUT CSRP1 was transiently cotransfected with GATA4 along with NPPA, VEGF, and NOS3-luciferase promoters respectively in HEK293 cells. Media was changed 3 h post transfection and cells were harvested for luciferase assay after 36 h. Relative luciferase activities were presented as fold changes. WT CSRP1 snergetically activates the promoters when cotransfected with GATA4, however, MUT CSRP1 abolished this synergy. The data represent the means of 3 independent experiments done in duplicates and the values are ± SE. P-value was assessed used Students' T-test. Significance $p<0.01$ is indicated by an (*) while $p<0.05$ is indicated by an (*); significance of synergy for WT is tested relative to the sum of individual activations, while that of mutant is tested relative to synergy. The triangle represents an increasing dose of the WT and MUT CSRP1 (400 ng and 600 ng respectively) and GATA4 (200 and 400 ng).

6. TBX5 is a novel partner for CSRP1

The T-Box transcription factor, TBX5, is involved in vertebrate cardiac and limb development and mutation in this gene cause amongst others cardiac septal malformation, similar to the ones observed in our family. Moreover, it was shown that TBX5 interacts with the chicken and zebrafish LMP4, a member of PDZ-LIM proteins(Camarata et al., 2006). Thus, we were interested in assessing a potential interaction between CSRP1 and TBX5. Co-immunoprecipitation results revealed that TBX5 is a strong physical partner of CSRP1 (figure 25A). Interestingly, the results of co-immunoprecipitation assays of HEK293cells overexpressing TBX5 and the variant revealed that the mutation decreased the interaction by up to 55.3%. We also assessed the functional interaction between CSRP1 and TBX5 by transient co-transfection assays in HEK293. The results showed a synergistic activation of these promoters reaching up to 22 folds (figure 25B-D); however, this synergy was completely abolished by the mutation in CSRP1.




B
NPPA Promoter
NPPA
luc
Fold Activation
25
20
15
10
5
0
**
**
WT CSRP1
MUT CSRP1
TBX5


C
VEGF Promoter
VEGF
luc
Fold Activation
25
20
15
10
5
0
**
**
WT CSRP1
MUT CSRP1
TBX5

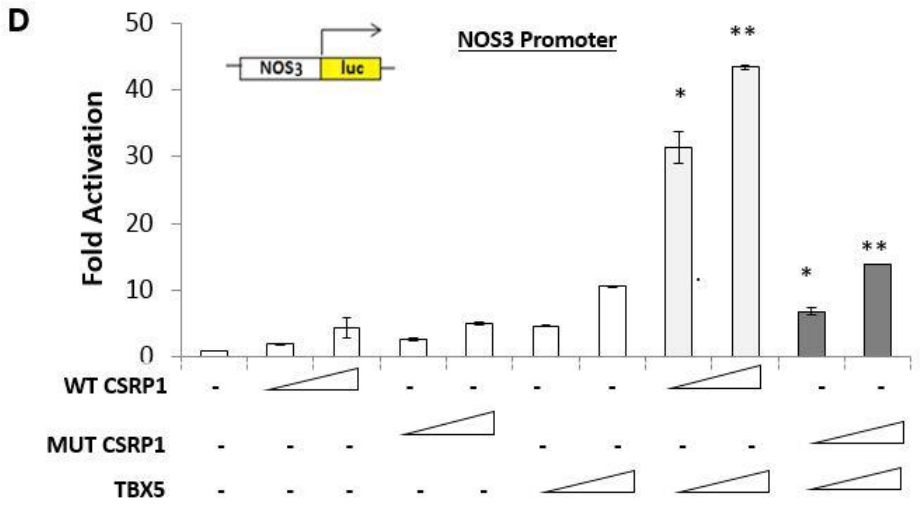


Figure 25. The effect of p.E154Vfs*99 mutation on the physical and functional interaction between CSRP1 and TBX5. (A) Physical interaction between HA-tagged TBX5 and Flag-tagged CSRP1 (WT and MUT) is demonstrated in the lanes of the right panel The amount of the proteins used for immunoprecipitation was ten times that used in western blot. Nuclear lysates of CSRP1 protein was immunoprecipitated with anti-HA antibody and CSRP1 proteins were detected using anti-Flag antibody. After membrane stripping, subsequent western blot analysis using anti-HA was performed to detect TBX5 protein. **(B, C, D)** Transcriptional activity of WT and/or MUT CSRP1 along with TBX5. WT or MUT CSRP1 was transiently cotransfected with TBX5 along with NPPA, VEGF, and NOS3-luciferase promoters respectively in HEK293 cells. Media was changed 3 hr post transfection and cells were harvested for luciferase assay after 36 hr. Relative luciferase activities were presented as fold changes. The data represent the means of 3 independent experiments done in duplicates and the values are ± SE. P-value was assessed used Students' T-test. Significance $p<0.01$ is indicated by (*), while $p<0.05$ is indicated by (*); significance of synergy for WT is tested relative to the sum of individual activations, while that of mutant is tested relative to synergy The triangle represents an increasing dose of the WT and MUT CSRP1 (400 ng and 600 ng respectively) and TBX5 (200 and 400 ng).

7. Exome Sequencing: A novel TRPS1 Variant

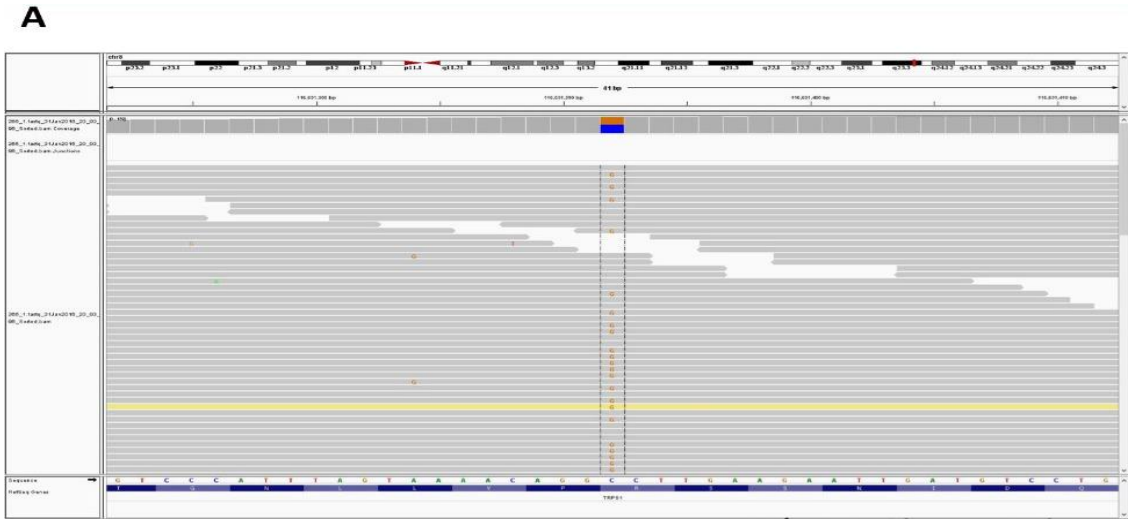
Since we did not find variations in CSRP1 partners including *SRF*, *GATA4*, and *TBX5* to explain the genotype of the phenotypically normal individual II.8 (Table 11), we decided to carry on whole-exome sequencing (WES) on selected members of the family (II.8, II.9, III.1, III.2, III.8, III.9, III.10, III.11, and III.12). We filtered the

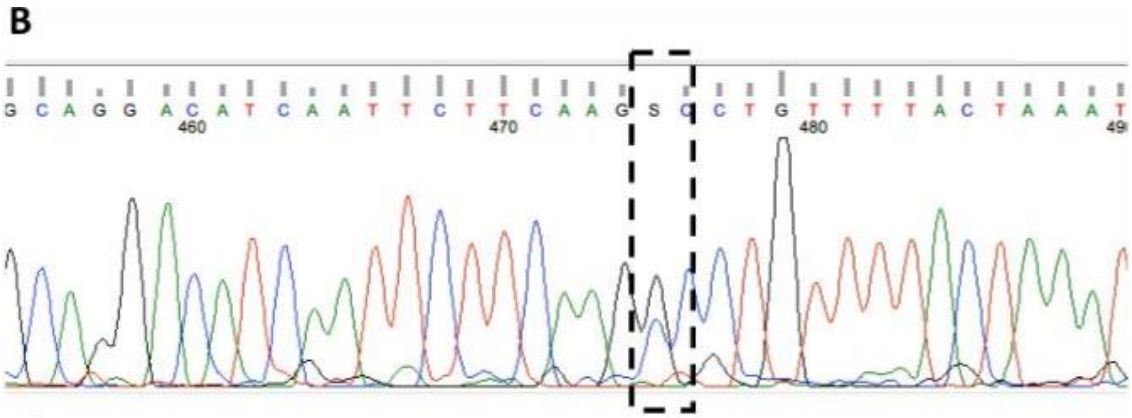
common inherited variants amongst affected individuals with the cardiac and/or polydactyly phenotype (Table is not shown in the thesis). We did not identify any rare damaging variants in polydactyly candidate genes to explain the limb phenotype in this family. We also did not identify similar variants in other genes that segregate with the polydactyly phenotype. We then interrogated the digenic hypothesis for CHD inheritance in the family by identifying rare damaging variants inherited from the mother (II.9) to the affected CHD probands. Using this approach, we identified a non-synonymous mutation in exon 8 of the *TRPS1* gene [NM_014112.4:c.933G>C] (Table 9). This variations leads to a missense p.R311S mutation (ENST00000519076) at the level of the protein (figure 26A-C). *In silico* analysis shows that the p.R2311S variant could have a probably damaging or neutral effect on the protein function, depending on the software used (SIFT, Polyphen2, Provean, MutationTaster). The variant is not found in the gnomAD database though highly covered both in the exome and genome databases (AF=0). We thus hypothesized that a combinatorial role for the two variants in *CSRP1* and *TRPS1* could explain the CHD phenotype.

8. Role of the digenic CSRP1/TRPS1 variants

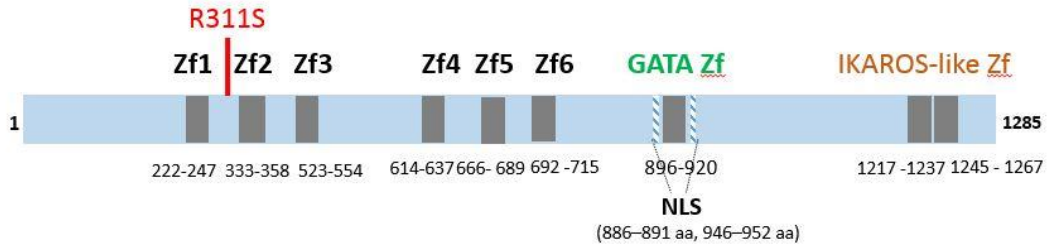
In order to assess the impact of the p.R311S variant on the functional properties of TRPS1, site-directed mutagenesis was carried out on the coding region of the human TRPS1 cDNA sub-cloned into an HA-tagged plasmid. Both plasmids were sequenced before transiently expressing them into Hela and HEK293 cells. Both the wild type and mutated proteins were equally produced as assessed by western blot analysis of nuclear proteins extracted from these cells (figure 26D). Cellular localization was assessed in Hela cells which showed that both the wild type and mutated proteins are present in

both the nuclei and cytoplasm of the transfected cells with no substantial differences (figure 26E). Functional luciferase assays on the three cardiac enriched promoters did not show any significant transcriptional activity (data not shown), prompting us to study the direct interaction between CSRP1 and TRPS1. Co-immunoprecipitation assays from HEK293 cells overexpressing both proteins show a relatively stable interaction between the two (figure 26F). Both mutations in CSRP1 and TRPS1 totally inhibit the interaction with the obligate partner (figure 26G, and data not shown), proving that both variants are deleterious.

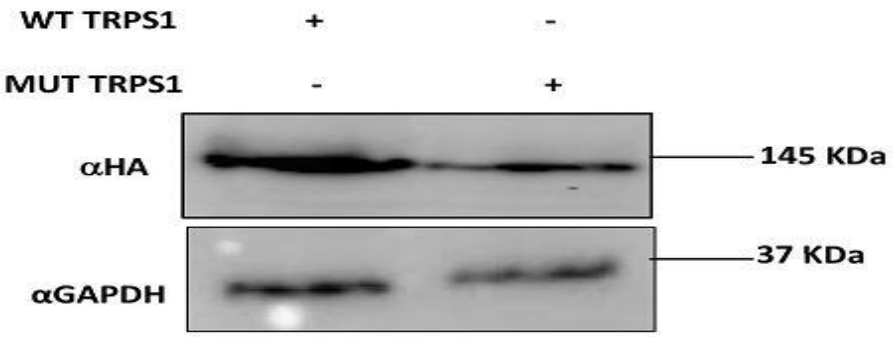




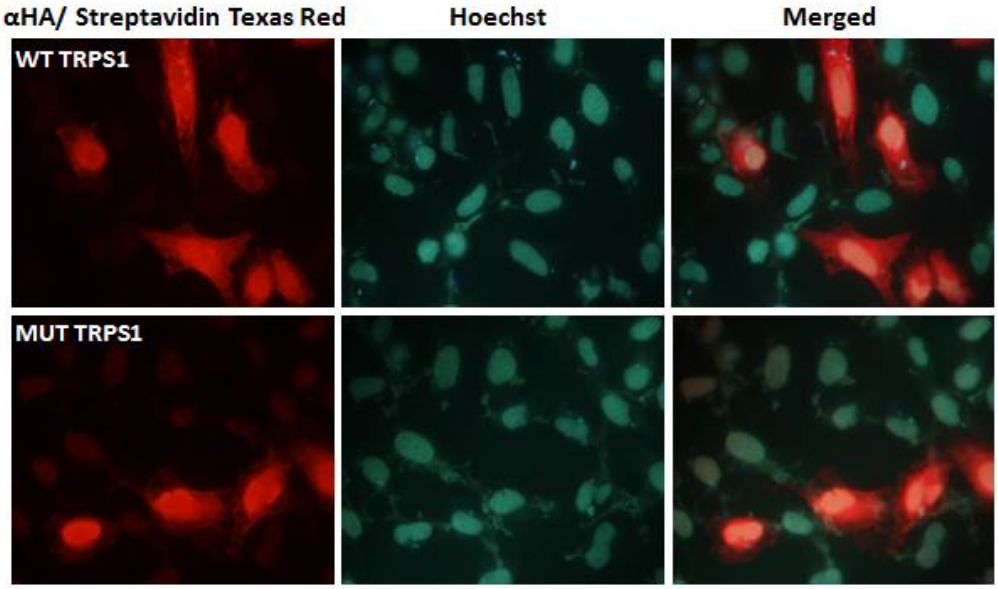
C


R311S
Zf1
Zf2
Zf3
Zf4
Zf5
Zf6
GATA Zf
IKAROS-like Zf
1
1285
222-247
333-358
523-554
614-637
666- 689
692 -715
896-920
1217 -1237
1245 - 1267
NLS
(886–891 aa, 946–952 aa)

D


WT TRPS1
+
-
MUT TRPS1
-
+
αHA
145 KDa
αGAPDH
37 KDa

E.


αHA/ Streptavidin Texas Red
Hoechst
Merged
WT TRPS1
MUT TRPS1

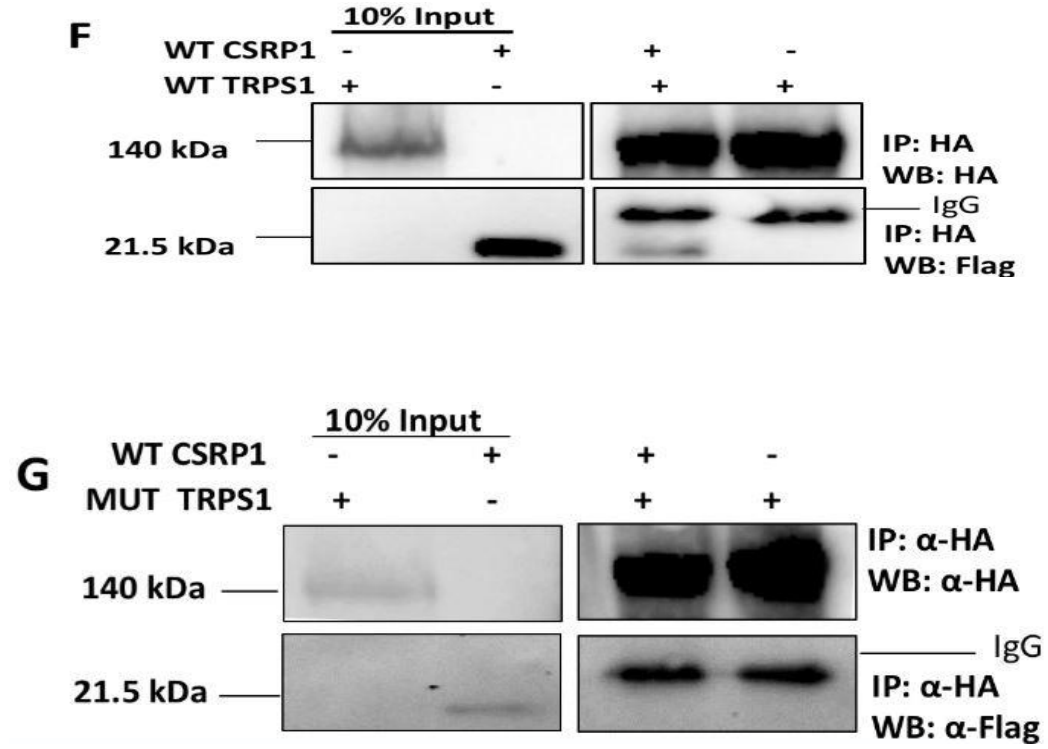


Figure 26. The physical interaction between TRPS1 and CSRP1. (A) Whole-exome sequencing shows a point mutation [NM_014112.4:c.933G>C] in Trsp1 gene. (B) Sanger sequencing of *TRPS1* gene confirms the point mutation G>C in exon 8. **(C)** Schematic representation of the TRPS1 protein with the different domains (Xf: Zinc finger. NLS: Nuclear Localization Signal), and the position of the mutation. **(D)** Nuclear extracts from transfected HEK293 cells with either WT TRPS1 or MUT TRPS1 were resolved on an SDS-PAGE. Immunoblotting using anti-HA antibody showed equal amounts of expressed proteins at 20 µg. Anti-GAPDH was used as a loading control. (**E**) Immunofluorescence of Hela cells transfected with WT TRPS1 and MUT TRPS1 plasmids. The localization of TRPS1 was visualized using anti-HA antibody followed by biotinylated anti-Rabbit antibody and then Streptavidin Texas Red for both WT TRPS1 and MUT TRPS1. Nuclei were stained blue with the Hoechst 33342 dye. TRPS1 (WT or MUT) showed cytoplasmic and nuclear localization (Red color). **(F)** Physical interaction between HA-tagged TRPS1 and Flag-tagged CSRP1 is demonstrated in the lanes of the right panel**.** The amount of the proteins used for immunoprecipitation was ten times that used in western blot. Nuclear lysates of CSRP1 protein was immunoprecipitated with anti-HA antibody and WT CSRP1 protein was detected using anti-Flag antibody. **(G)** This interaction is lost when using the MUT TRPS1 protein instead of the WT.

E. Discussion

We used both targeted exome and whole exome sequencing to unravel the genetic factors responsible for both cardiac malformations and polydactyly in a large Lebanese family with high consanguinity. We showed that despite the high

consanguinity between members of the family, there were no homozygous mutations that could account for either or both phenotypes in the annotated genes. In addition, we found that both phenotypes are not linked, and that the cardiac phenotype is associated with a novel mutation in the gene encoding CSRP1, the first LIM domain protein implicated in CHD. In addition, we interrogated the penetrance of this mutation to the potential modifiers that could impact its complete penetrance, and revealed a novel di-genic interaction involving the zinc finger encoding gene *TRPS1*.

1. CSRP1: A specific integrator of cardiac-enriched transcription factors

To the best of our knowledge, this is the first *CSRP1* variant identified in CHD patients or in any other cardiovascular diseases. The p.E154Vfs*99 mutation did not affect the cellular localization of the protein; however, it affected the structure and function of CSRP1 protein. The disruption of the protein physical interactions with its partners and the inhibition of the transcriptional activity on several cardiac enriched promoters suggest the severity of this mutation which affects the cardiac phenotypic outcome in the family. Our *in silico* analysis showed that this mutation is disrupting the second LIM domain at the C-terminal of the protein. The LIM domain has been shown to be implicated in protein-protein interactions(Arber et al., 1994; Schmeichel & Beckerle, 1994). Since the mutation lies in an essential domain of the protein, we suggested a conformational, structural and functional change at the protein level. This was corroborated by the obstruction of the physical interaction between CSRP1 and its partner SRF, a transcription factor mandatory for the appearance of cardiac mesoderm during embryonic mouse development and an essential partner of cardiac GATA4(Small & Krieg, 2003; Wang et al., 2002). Similarly, the p.E154Vfs*99 has

abrogated the interaction between CSRP1 and GATA4 which is also a vital transcription factor in the early and late heart development such as valve formation and cardiac septation(Jeffery D. Molkentin et al., 1997). Dominant *GATA4* mutations cause severe congenital heart disease including atrial and atrioventricular septal defects (ASDs and AVSDs) as well as Tetralogy of Fallot(Garg et al., 2003b; Georges Nemer et al., 2006). This role of GATA4 in multiple forms of congenital heart disease is reminiscent of our findings for CSRP1 in this particular family, and could be explained by the broad yet timely expression pattern of the proteins during heart development. Indeed, we showed that CSRP1 is expressed in all cardiac compartments, and it was previously shown that a CRE-driven CSRP1 enhancer is highly expressed in the outflow tract, and both in the mesenchymal cells as well as cardiomyocytes supporting a role for CSRP1 in cardiac development(Snider et al., 2008). Interestingly, we showed that the mutation in CSRP1 did not affect its interaction with GATA5 and/or GATA6 suggesting different interfaces with different outcomes on different promoters. Additionally, we documented a novel interaction of CSRP1 with TBX5, a member of the T-box family, implicated in the Holt-Oram syndrome. This interaction is not novel between the two classes of protein, since it was previously shown that the LIM4 and pdlim7 proteins regulate cellular localization of TBX5 during pectoral and heart development(Camarata et al., 2006; Camarata, Krcmery, et al., 2010; Camarata, Snyder, et al., 2010). In our case, the interaction is functional between the two protein resulting in a synergistical activation of downstream target genes like NPPA. Although our results show that CSRP1 can coordinate protein partners through its LIM domain to form a robust network of transcriptional activators, it is still possible to speculate over the contribution of the

LIM DNA binding-domain on its own on specific promoters (Kadrmas & Beckerle, 2004).

Despite the potential link that this interaction could suggest over the polydactyly phenotype observed in members of the family, the mutation in CSRP1 does not segregate with the limb phenotype and thus exclude any role this mutation could have on the limb defects. This dichotomy in the genetic inheritance of two different phenotypes is not however novel in the case of CHD and limb defects. In fact besides syndromic cases caused by monogenic mutations like *TBX1,3* and *5, SALL4*, and *EVC*, there are no published studies on variants that cause only cardiac and limb defects whether monogenic or multigenic (Hills, Kochilas, Schimmenti, & Moller, 2011; Jürgen Kohlhase et al., 2002; Packham & Brook, 2003).

2. CSRP1/TRPS1: A new digenic paradigm in CHD

Since CHD is a multifactorial disease in that it frequently reveals variable penetrance, genetic heterogeneity, and variable expressivity, it was essential to unravel other partners for CSRP1 that could account for the partial penetrance problem we faced in this particular family. Since the *CSRP1* variant is inherited from the father's side (II-8) who yet has a normal cardiac phenotype but has a limb defect, it was instrumental to go for whole-exome sequencing since cardiac targeted sequencing was not enough to explain the cardiac phenotypes in probands and we could not find a variation that explains the limb phenotype. The extracted data failed to show any monoallelic variation in genes previously shown to be linked to polydactyly(Biesecker, 2011) nor to yield a common variant between the 7 members with the same phenotype included in the screening. In contrast, the WES data yielded multiple variants inherited from the

mother (II-9) and only present in the cardiac-affected probands. The *TRPS1* variant (p.R311S) stands alone among a short list of missense variants (Supplementary table 2, not shown in the thesis), since it was novel, absent from the gnomad database, as well from 200 Lebanese exomes. The *in silico* analysis shows that this variant would have a moderate effect on the function of the protein, and we thus hypothesized that alone it would not affect cardiac development, thus explaining the normal phenotype of the mother. Since probands III-8,-10 and -12 who carry the *CSRP1* and *TRPS1* variants expressed a cardiac phenotype, we hypothesized that the *TRPS1* variant has no effect unless it is expressed with the *CSRP1* variant. Mutations or deletions in *TRPS1* give rise to Tricho-rhino-phalangeal syndrome (TRPS), and a relatively high proportion of patients with TRPS exhibit CHD, ranging from minor to severe anomalies(Maas et al., 2015a; Verheij et al., 2009b). Although previous reports did not describe Trps1 expression in the mammalian heart, Trps1 was recently shown to be expressed in a restricted region within the cardiac cushion of OFT and developing valves(Nomir et al., 2016). This could partially explain why patients with *TRPS1* mutations show a broad range of congenital cardiac defects. In our case, we establish a direct physical interaction between the two proteins that was affected by the CSRP1 variant. However, we suspect a much more complicated functional regulation since TRPS1 acts as a repressor and CSRP1 as a weak activator of downstream target genes. We hypothesize that the transcriptomal assembly of CSRP1 and TRPS1 co-factors over cardiac promoters, would be largely affected by the mutations, and cause the observed phenotypes in the family members along the broad spectrum of phenotypes associated to CHD.

Although efficient strategies such as whole exome sequencing potentially contribute to the understanding of rare human diseases and allows the detection of multiple rare variants, they are still short of elucidating the network of such genes that are involved in CHD(Postma et al., 2015). Thus, genome-wide association studies (GWAS) should be combined with WES or even WGS since they have provided evidence that common genetic variation can influence the risk of certain types of CHD (detect somatic mutations and noncoding sequences). Many whole-genome CNV screening studies indeed have revealed that a significant number of CHD patients have pathogenic CNVs. The highest frequency of pathogenic CNVs are found in patients who have both CHD and extra-cardiac anomalies which is similar to our case(Andersen, Troelsen, & Larsen, 2014).

3. Conclusion

Dissecting phenotypes and establishing direct genotype/phenotype interaction is a must in any strategical approach in CHD. We have shown that despite the high consanguinity within one family, there are no homozygous mutations that could explain the associated cardiac and limb defects. In contrast, a digenic model of inheritance is proposed to explain the cardiac phenotype.

CHAPTER VI

GENERAL DISCUSSION

CHD is described as developmental abnormalities triggered by mutations in genes and pathways which are involved in the spatiotemporal proliferation and differentiation of cardiovascular cells. CHD can occur independently or in conjunction with other anomalies in different organs (Marian, 2017). In this work, targeted sequencing enabled us to identify a novel heterozygote duplication of 14 nucleotides leading to a frameshift mutation (p.E154Vfs*99) in *CSRP1* gene in a Lebanese Family with both Cardiac-limb deformities. However, genotype-phenotype correlations revealed that this mutation segregates better with the cardiac phenotype. On the other hand, we revealed by WES a novel missense mutation in *TRPS1* (p.R311S) inherited from the normal mother and only present in patients with the cardiac phenotype. Through *in vitro* approaches, we showed that p.E154Vfs*99 mutation did not affect the cellular localization of the protein; however, it affected the structure and function of CSRP1 protein. The disruption of physical interactions of CSRP1 with its partners and the inhibition of the transcriptional activity on several cardiac enriched promoters including NPPA imply a severe effect of this mutation on cardiac development. Thus, we speculate that this variant might be involved in the etiology of CHD in this family. To the best of our knowledge, this is the first CSRP1 variant identified in CHD patients.

Because of the high degree of consanguinity between family members, one could expect a homozygous mutation to be present in the probands. Many studies have been done on the relationship between consanguinity and CHD in which VSD and ASD are the most mentioned defects(Bittles, 2011). However, our data reveal that although

VSD and ASD are common in this family, still there is not any homozygous mutation accounting for either CHD or polydactyly. This is supported by the fact that many of the most common CHDs are genetically heterogeneous, whether detected as isolated anomalies or accompanied by other heart deformities. Till now, it is unclear what is the nature and significance of the link between consanguinity and CHD (Bittles, 2011).

A. p.E154Vfs*99 CSRP1: a deleterious variant

Our *in silico* analysis supported by Polyphen-2 prediction tool revealed that p.E154Vfs*99 variant is damaging and thus disease causing. Specifically, p.E154Vfs*99 lies at C-T LIM domain (LIM2), an essential domain of the CSRP1. It has been previously shown that the LIM domain is implicated in protein-protein interactions(Arber et al., 1994; Schmeichel & Beckerle, 1998). Consequently, we suggested a conformational, structural and functional modification at the protein level. This was supported by the disruption of the physical interaction between CSRP1 and its partner SRF (MADS box protein)(David J. McCulley & Black, 2012), a transcription factor required for the appearance of cardiac mesoderm during embryonic mouse development and an essential partner of cardiac GATA4 (Sepulveda et al., 2002). Similarly p.E154Vfs*99 has abrogated the interaction between CSRP1 and GATA4 which is also a vital transcription factor in the early and late heart development such as valve formation and cardiac septation(Peterkin et al., 2005; L. Yang, 2011). *Gata4* knockout mice display severe defects in cardiac morphogenesis which causes lethality at E8(Lentjes et al., 2016). Dominant GATA4 mutations cause severe congenital heart disease (CHD) including atrial and atrioventricular septal defects (ASDs and AVSDs). It was recently demonstrated that that GATA4 is needed in “SHF” progenitor cells

which are responsible for cardiac septation(Zhou et al., 2017) . This role of GATA4 in multiple forms of congenital heart disease is reminiscent of our findings for CSRP1 in this particular family, and could be explained by the broad yet timely expression pattern of the proteins during heart development.

Interestingly, p.E154Vfs*99 mutation did not affect the interaction of CSRP1 with GATA6 while WT CSRP1 did not interact with GATA5. Thus, we suggest that GATA4,5, and 6 have variable interfaces with different outcomes on different promoters and this may be due to some differences in their binding site preferences(Sakai, Nakagawa, Sato, & Maeda, 1998).

It was previously reported that GATA4 functions as an important partner for the T-box transcription factor TBX5 which is crucial for both cardiac and limb development(Peterkin et al., 2005). Specifically, TBX5 is crucial for cardiac septation. *Gata4* and *Tbx5* double heterozygous mice die at E15.5 due to VSDs and ASDs(Maitra et al., 2009) . The cardiac malformations seen in humans when GATA4 is disrupted are similar to those seen with Holt-Oram syndrome mutations in TBX5(Peterkin et al., 2005) . Our results on the interaction between CSRP1 and TBX5 along this pathway support a role of CSRP1 in cardiac septation. Nonetheless, this interaction is not novel between the two classes of proteins, since it was previously shown that the LIM4 and pdlim7 proteins regulate cellular localization of TBX5 during pectoral and heart development (Camarata et al., 2006, 2006; Camarata, Snyder, et al., 2010). As the WT CSRP1 physically interacted with TBX5, CSRP1 variant hindered this interaction. We thus suggest that the abrogation of interaction between CSRP1 and TBX5 is disease causing.

Still, it is essential to determine the domains of interaction between CSRP1 and each partner through generating CSRP1 proteins with variable domain deletions in order to predict benign from deleterious variations in the different proteins.

B. CSRP1 expression in the heart: A phenotype-Genotype Correlation

Despite the deleterious effect of the mutation on the protein function, there were no supporting data of the expression of CSRP1 protein in the heart during early phases of development. It was previously shown that a CRE-driven CSRP1 enhancer is highly expressed in the outflow tract, and both in the mesenchymal cells as well as cardiomyocytes supporting a role for CSRP1 in cardiac development(Snider et al., 2008). Those data corroborate with what we have found in cardiomyocytes for CSRP1 expression in which we showed by immunohistochemistry that CSRP1 is expressed throughout the cardiac compartments. Moreover, Chang et al. showed that CRP2 was found to be localized to the nucleus at early stages of smooth muscle differentiation(Chang et al., 2003b). Similarly, our results for CSRP1 expression *in vivo* show that we have a higher expression of the protein in the nucleus of cardiomyocytes as compared to the cytoplasm at different embryonic stages. Specifically, a nuclear expression of CSRP1 in the atrial septal region of the mouse heart (E12.5) appears proposing a possible role of CSRP1 in the septal formation but not in the valves (supplementary figure 1). Of note the absence of the expression of the protein in the endocardial cells and the progressive increase in the cytoplasmic localization of the protein whereby the newborn heart mice (one week old) have mainly cytoplasmic CSRP1 (supplementary figure 2). This explains that CSRP1 have different functions throughout development possibly ranging from the co-activation of transcription in

proliferating and early differentiating cardiac progenitor cells through its interaction with essential cardiac transcription factors SRF, GATA4 and TBX5, to a cyto-architectural role in mature cardiomyocytes. The latter is due to the fact that CSRP1 interacts with α-actinin and zyxin which are actin crosslinking proteins(Louis et al., 1997; Pomiès et al., 1997). Unfortunately, it still unclear how CSRP1 shuttles between the cytoplasm and the nucleus, and thus signaling pathways that control this dual localization of the protein should be elucidated. The unraveling of the effects of post translational modifications such as phosphorylation of CSRP1 on its subcellular localization is a future goal, if CSRP1 was found to be implicated in adult cardiac diseases like cardiomyopathies.

In addition to the expression of CSRP1 throughout the heart, the multiple interactions of CSRP1 with various partners on different cardiac-enriched promoters (NPPA, VEGF, and NOS3) supports its possible involvement of CSRP1 in several genetics pathways that regulate normal heart development and might be linked to the variable clinical expression of CHD in this family.

Unfortunately, we failed to map an essential CSRP1 binding domain on NPPA promoter by luciferase assay through sequential NPPA promoter deletions (data not shown), in which we yielded similar patterns of activation for all the deleted domains on the promoter. We almost expected these results since although the zinc binding carboxy-terminal module of LIM domain in CRP proteins and all LIM domains is structurally comparable to that observed for DNA-binding domains of GATA-1 and glucocorticoid receptor, there is yet no evidence for sequence specific interactions between LIM domains and DNA sequences despite the structural similarity of LIM

domains to DNA-binding motifs. In contrast, many LIM domain proteins have been reported to mediate specific interactions with other protein partners(Yao et al., 1999). Thus, it is highly recommended to detect CSRP1 downstream target genes and a specific binding site by ChIP-seq technique which would be further confirmed by electrophoretic mobility shift assay (EMSA).

C. Incomplete Penetrance: Exome and Genome sequencing should solve it

CHD is a multifactorial disease in that it frequently reveals variable penetrance, genetic heterogeneity, and variable expressivity. In addition, CHD is influenced by both genetic and environmental factors (Ai et al., 2010). Thus, identical CHD subtypes can be triggered by various mutations. Moreover, even if different patients belong to the same family, the same mutation can still cause variable phenotypes (Zhang et al., 2016). Incomplete penetrance is one of the major paradigms that are not answered in genetic studies. In our case, it was clear that the CSRP1 variant is inherited from the father's side II (-8) who yet has a normal cardiac phenotype but has a limb defect, it was thus essential to explore whole-exome sequencing (WES) since targeted sequencing was not enough to explain the cardiac phenotypes in probands and we could not find a variation that explains polydactyly. WES has yielded many variants among which we concentrated on the ones inherited from the mother and are present in the affected probands only. A novel missense mutation (p.R311S) in *TRPS1* is the one that caught our attention in this case after filtering the benign variants. Although our *in silico* analysis using Polyphen-2 predicted that this variant is deleterious and probably damaging, the mother (II-9) has the p.R311S TRPS1 variant and reveals a normal cardiac and limb phenotype. Since probands III-8,-10 and -12 who carry the CSRP1 and

TRPS1 variants expressed a cardiac phenotype, we hypothesized that TRPS1 variant has no effect unless it is expressed with CSRP1 variant. Mutations or deletions in TRPS1 give rise to Tricho-rhino-phalangeal syndrome (TRPS), and a relatively high proportion of patients with TRPS exhibit CHD, ranging from minor to severe anomalies(Maas et al., 2015a; Verheij et al., 2009b). Although previous reports did not describe Trps1 expression in the heart, Trps1 was recently shown to be expressed in a restricted region within the cardiac cushion of OFT and developing valves(Nomir et al., 2016). This could partially explain why patients with TRPS show a broad range of congenital cardiac defects(Maas et al., 2015a). We thus hypothesize that our TRPS1 variant could affect heart development. We first assessed TRPS1 transcriptional activity on NPPA promoter, and there was no significant change in transcriptional activation (supplementary figure 3). We showed instead a direct physical interaction with CSRP1 that was affected by both variants. However, the pathways that regulate TRSP1 and CSRP1 expression in the heart are not clear and should be elucidated. We thus suspect a much more complicated functional regulation since TRPS1 acts as a repressor and CSRP1 as a weak activator of downstream target genes. We hypothesize that the transcriptomal assembly of CSRP1 and TRPS1 co-factors over cardiac promoters, would be largely affected by the mutations and affect the observed phenotypes in the family members along the broad spectrum of phenotypes associated to CHD. We have proposed a digenic mode of inheritance to explain the cardiac phenotype in our family. A digenic mode of inheritance involving BBS genes and MKKS gene, was previously reported in patients with Bardet-Biedl syndrome (BBS) which is a genetically heterogeneous disorder characterized by many features including polydactyly, obesity, mental retardations and retinal dystrophy. In addition, patients with BBS have the risk

of developing other features such as congenital heart disease and hypertension(Fauser, Munz, & Besch, 2003). However, we failed to link a variant to polydactyly in this family and the digenic model of inheritance we proposed is responsible for one phenotype in this family which is CHD.

D. Monozygotic Twins: the failure of the genetic code

Although twin studies provide good models to investigate the etiology of CHD and reveal the most rational results for sorting out genetic factors from environmental effects, our monozygotic (MZ) twins has variable phenotypes. One of which has severe CHD condition (ASD) and the other has a mild cardiac condition (PFO) that has no threat on one's life. Yet, they both carried the CSRP1 and TRPS1 variants. Although in general, the monozygotic (MZ) twins should reveal the same phenotype since they have the same genetic background, they are raised separately under the influence of different environmental factors, and thus their cardiac defects are different in many families (Xike Wang et al., 2014; Zhang et al., 2016). Chaiyasap et al. recently identified a pair of MZ twins with trisomy 21 but discordant for a ventricular septal defect and epilepsy. However, no discordant DNA variants were detected in both twins after Sanger sequencing. In addition, Zhang R. et al. performed WES for nine MZ pair twins, five of which are discordant for CHD phenotype. Although WES analysis revealed a *de novo* variation in TMPRSS13 in one of the CHD-affected twins; however, Sanger sequencing failed to confirm a variation in the affected twin. As others, we were unable to detect a discordant variable in these two MZ twins through WES and targeted sequencing. One explanation for the different phenotypes in the twin could be due to post-twinning mutations that might include single nucleotide mutations, CNV, or epigenetic

changes(Chaiyasap et al., 2014; Zhang et al., 2016). Another explanation for discordant phenotypes in twins is that individuals with less severe phenotypes, for example in our case the twin with PFO, could harbor mosaic genomic alterations that did not reach the appropriate level for clinical expression (Chaiyasap et al., 2014). This highlights a limitation of WES in detecting causative variants for discordant phenotypes in MZ twins.

E. Polydactyly: A distinct phenotype requiring different approaches

Although efficient strategies such as WES potentially contributes to the understanding of rare human diseases and allows the detection of multiple rare variants, it can only detect mutations in the protein coding genes which constitutes 1% of the human genome(Choi et al., 2009). Thus, WES will not completely determine the causes of CHD at least not with the predominant paradigm analysis that WES will always identify disease causing variants(Postma et al., 2015) . This might be one explanation for not detecting a polydactyly variant. Thus, linkage studies are highly recommended in our case since they have provided evidence that common genetic variations can influence the risk of certain types of CHD (detect somatic mutations and noncoding sequences). In addition, this type of high-throughput sequencing which can be applied on different tissues, developmental stages, and pathological conditions have the ability to reveal potential new regulators of heart development and disease. In addition, Many whole-genome CNV screening studies have been published which comprise more than 5000 patients(Anderson et al., 2014). These studies reveal that a significant number of CHD patients have pathogenic CNVs. The highest frequency of pathogenic CNVs is found in patients who have both CHD and extra-cardiac anomalies which is similar to

our case. In contrast, the frequency of pathogenic CNVs is significantly lower among patients with an isolated CHD in which only one gene is responsible for CHD in those patients. Overall, since most CNV loci harbor multiple genes, it is hard to identify the precise cause of the disease and unravel its pathophysiological mechanism(Postma et al., 2015). It is also well established that through WES or WGS not all regions of the exome or genome can be covered to an appropriate depth. Thus, if a causal variant is located in an uncovered or poorly covered region, it will be missed(Blue & Winlaw, 2015). That could be another possible explanation of why we did not find a causative polydactyly variant. In addition, we are not the first to fail in unraveling a variant for a certain disease. For example, Martin *et al.* efforts in using WES were not successful in determining a causal variant in a multiplex family with Bicuspid aortic valve (BAV) in addition to other cardiovascular deficiencies(Shiaulou Yuan, Samir Zaidi, 2014).

F. Limitations and drawbacks

Our study identified two novel mutations in CSRP1 and TRPS1 respectively that likely contributed to the CHD in this family. This finding is the first to underscore the pathogenic correlation CSRP1 mutation and CHD. Our study would significantly contribute towards understanding CHD and development pathways involved. However, the exact role of CSRP1 and TRPS1 in CHD and cardiac formation remains to be answered in the future.

As all research projects, our project has some limitations and drawbacks. The first limitation is that the part related to sequencing was time consuming because samples collected from the Lebanese family had to be sent abroad to be sequenced and analyzed. Although we performed WES, an advanced technique used in case of

heterogeneous diseases, we were not able to link a variant to polydactyly. Probably, sequencing of the remaining members would permit us to detect a causative variant for polydactyly. Although we used WES, linkage association studies are highly recommended. Nevertheless, they are still considered expensive especially after recruiting more members of the family.

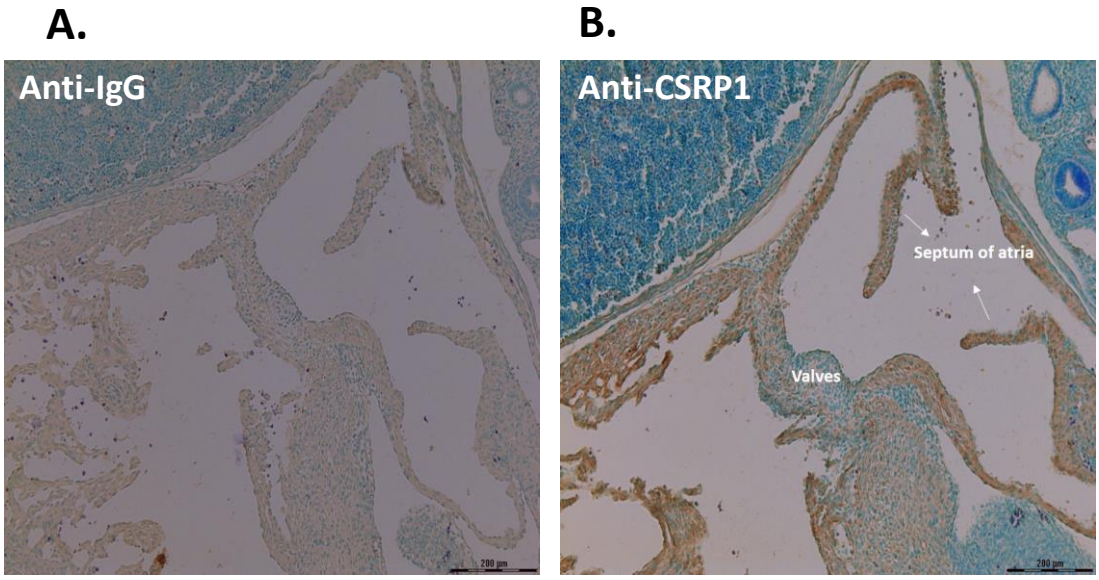
Although we have assessed the physical and functional interaction between CSRP1 and its protein partners, all our experiments were performed *in vitro* using HEK293 (embryonic) and Hela (cancerous) cells as models. Those cells absolutely have differential signaling machinery as compared to cardiomyocytes, and thus could not be completely representative of what is happening *in vivo*. Thus, *CSRP1* Knock-out mice would provide a better insight into the physiological role of *CSRP1* gene in humans and specifically in cardiac development and congenital heart disease. It would be of valuable interest to analyze the *CSRP1* knock-out mice and look at the cardiac phenotype of these animals especially that those animals are viable after the *CSRP1* knock-out. On the other hand, this subtle mutation can be mimicked in a mouse model using gene targeting. Instead of disrupting *CSRP1* gene, as in most knock-out mice, homologous recombination is employed to swap the normal copy of the exon with p.E154Vfs*99 mutation. As long as a similar mutation can be reproduced in the mouse protein, then the amino acid insertion can be targeted into a gene of interest to replicate the human disease. The effects of the altered protein can then be studied in the animal model. Unfortunately, obtaining *CSRP1* knock-out mice and generating the mutation *in vivo* are difficult and long-delayed. Ultimately, generating double heterozygous mice with both *CSRP1* and *TRPS1* variants and analyzing potential cardiac phenotypes can help us understand more the cardiac phenotypic outcomes of our patients.

Although we have detected several potential nuclear partners of CSRP1 *in vitro*, this does not mean that those are the only proteins that associate with CSRP1 especially that C-terminal LIM domain was predicted to have at least one protein partner. More experiments should be conducted to detect other partners of CSRP1 in the nucleus and cytoplasm due to the dual expression of CSRP1 in the cell. In addition, since there is yet no evidence for sequence-specific interactions between LIM domain proteins and DNA, we will be the first to determine CSRP1 binding to nucleic acid through performing ChIP-seq. However, this technique is sensitive, expensive and time consuming especially that it needs further confirmation with Gel shift assay.

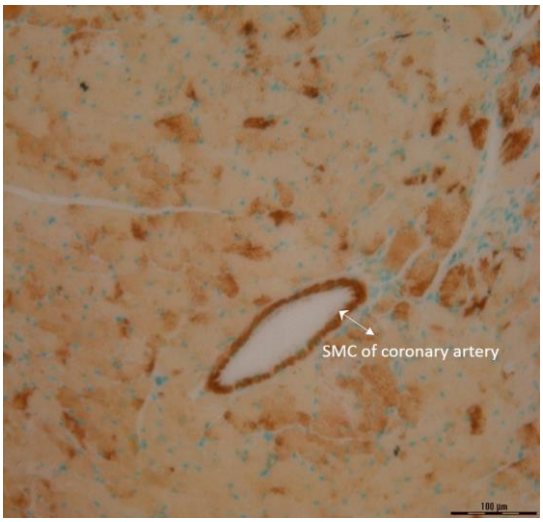
APPENDICES

APPENDIX A

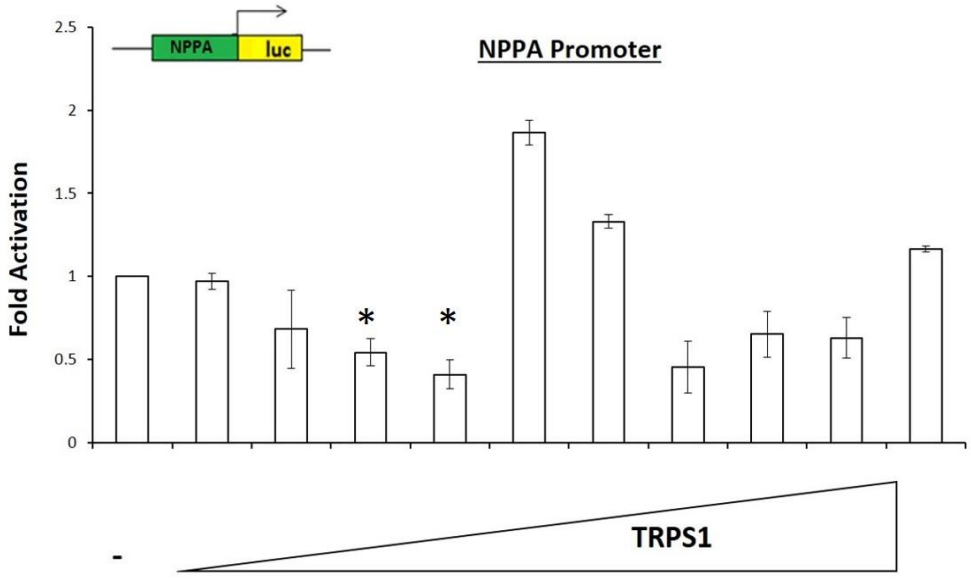
SUPPLEMENTARY FIGURES



Supplementary figure 1. The expression of CSRP1 at the atrial septal region of embryonic mouse heart (E12.5). A. Rabbit Anti-IgG (light brown color) is used as a negative control on a mouse section (E12.5) to show the specificity of anti-CSRP1**. B.** CSRP1 is highly expressed in the septal region (dark brown dots) of the cardiac septum of the atria, but not in the valves (light brown color). Nuclei were counter-stained with methyl green. Photos were taken at magnification 10x.



Supplementary figure 2.The expression of CSRP1 in an adult mouse heart. CSRP1 shows a higher expression in the cytoplasm (dark brown color) than in the nuclei (light blue dots) in the heart of an adult mouse. However, in smooth muscle cells (SMC) of the coronary artery, CSRP1 expression appears to be more nuclear (dark brown dots). Nuclei were counter-stained with methyl green. Photos were taken at magnification 10x.



Supplementary figure 3. Transcriptional activity of WT TRPS1 on NPPA-luciferase promoter. WT TRPS1 was transiently transfected along with 3.5 µg of NPPA-luciferase promoter in HEK293 cells. WT TRPS1 significantly represses NPPA promoter activity (*) as compared to the control (-).P-value was assessed used Students' T-test. Significance $p<0.05$ is indicated by (*); significance is tested relative to control. Relative luciferase activities were presented as fold changes. The data represent the means of 2 independent experiments done in duplicates and the values are ± SE. P-value was assessed using Students' T-test .Significance is tested relative to control (-). The triangle represents an increasing dose of the WT TRPS1 (100ng-1000ng).

APPENDIX B

SUPPLEMENTARY TABLES

Supplementary Table 1. List of 119 genes implicated in CHD or cardiac development as previously described (Abou Hassan et al., 2015)

ACTA2	ELN	ISL1	NOTCH1	SESN1
ACTC1	EP300	JAG1	NOTCH2	SHOC2
ACVR1	ESCO2	KCNJ2	NPHP3	SLC2A10
AHSA2	EVC	KIAA1841	NSD1	SOS1
ANKRD1	EVC2	KIF3C	OSR1	SRF
ASXL2	FBN1	KLF13	PAPOLG	TBX1
ATE1	FBN2	LBR	PCMTD2	TBX20
BCL11A	FGFR1	LEFTY2	PCSK5	TBX3
BCOR	FLNA	LRP2	PEX1	TBX5
BMP4	FOXC1	MAX	PEX13	TCF21
BMP7	FOXH1	MED13L	PITX2	TDGF1
BMPR2	FOXL2	MEF2A	PKD2	TFAP2B
C1orf106	GATA4	MEF2C	PLAGL1	TLL1
CCT4	GATA5	MGP	PPM1K	TWIST1
CFC1	GATA6	MID1	PPP3CA	UBR1
CHD7	GDF1	MKKs	PQBP1	USP34
CITED2	GJA1	MSX1	PROX1	VEGFA
CREBBP	GJA9	MSX2	PTPN11	VEGFC
CRELD1	GPC3	MYH11	RAB10	XPO1
CRX	HAND1	MYH6	RAB23	ZEB2
CSRP1	HAND2	NF1	RAI1	ZFPM1
CTNNA3	HES1	NFATC3	RAI2	ZFPM2
DHCR7	HES4	NFATC4	REL	ZIC3
DNAI1	HEY2	NIPBL	ROR2	ZNHIT3
DQ983818	HOXA1	NKX2-5	SALL1	
DVL1	ID2	NKX2-6	SALL4	
EHMT1	IGFBP4	NODAL	SEMA3E	

BIBLIOGRAPHY

Abou Hassan, O. K., Fahed, A. C., Batrawi, M., Arabi, M., Refaat, M. M., DePalma, S. R., … Nemer, G. M. (2015). NKX2-5 Mutations in an Inbred Consanguineous Population: Genetic and Phenotypic Diversity. *Scientific Reports*, *5*, 8848. https://doi.org/10.1038/srep08848

Abu-Issa, R., Smyth, G., Smoak, I., Yamamura, K., & Meyers, E. N. (2002). Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. *Development (Cambridge, England)*, *129*(19), 4613–4625.

Aburawi, E. H., Aburawi, H. E., Bagnall, K. M., & Bhuiyan, Z. A. (2015). Molecular insight into heart development and congenital heart disease: An update review from the Arab countries. *Trends in Cardiovascular Medicine*, *25*(4), 291–301. https://doi.org/10.1016/j.tcm.2014.11.007

Ahn, D., Kourakis, M. J., Rohde, L. A., Silver, L. M., & Ho, R. K. (2002). T-box gene tbx5 is essential for formation of the pectoral limb bud. *Nature*, *417*(February), 754–758. https://doi.org/10.1038/nature00814

Ai, J., Zhang, R., Li, Y., Pu, J., Lu, Y., Jiao, J., … Yang, B. (2010). Circulating microRNA-1 as a potential novel biomarker for acute myocardial infarction. *Biochemical and Biophysical Research Communications*, *391*(1), 73–77. https://doi.org/10.1016/j.bbrc.2009.11.005

Al-Baradie, R., Yamada, K., St Hilaire, C., Chan, W.-M., Andrews, C., McIntosh, N., … Engle, E. C. (2002). Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. *American Journal of Human Genetics*, *71*(5), 1195–1199. https://doi.org/10.1086/343821

Allanson, J. E., & Newbury-Ecob, R. A. (2003). Holt-Oram syndrome: Is there a "face"? *American Journal of Medical Genetics*, *118A*(4), 314–318. https://doi.org/10.1002/ajmg.a.10214

Andersen, T. A., Troelsen, K. de L. L., & Larsen, L. A. (2014). Of mice and men: molecular genetics of congenital heart disease. *Cellular and Molecular Life Sciences : CMLS*, *71*(8), 1327–52. https://doi.org/10.1007/s00018-013-1430-1

Anderson, R. H., Spicer, D. E., Brown, N. A., & Mohun, T. J. (2014). The development of septation in the four-chambered heart. *Anatomical Record*, *297*(8), 1414–1429. https://doi.org/10.1002/ar.22949

Arber, S., & Caroni, P. (1996). Specificity of single LIM motifs in targeting and LIM/LIM interactions in situ. *Genes and Development*, *10*(3), 289–300. https://doi.org/10.1101/gad.10.3.289

Arber, S., Halder, G., & Caroni, P. (1994). Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. *Cell*, *79*(2), 221–231. https://doi.org/10.1016/0092-8674(94)90192-9

Bang, O. Y., Lee, J., Ryoo, S., Kim, J., & Kim, W. (2015). Patent foramen ovale and stroke – Current status. *Journal of Stroke*, *17*(3), 229–237.

https://doi.org/10.1161/CIRCULATIONAHA.104.524371

Basson, C. T., Bachinsky, D. R., Lin, R. C., Levi, T., Elkins, J. A., Soults, J., … Seidman, C. E. (1997). Mutations in human cause limb and cardiac malformation in Holt-Oram syndrome. *Nature Genetics*, *15*(1), 30–35. https://doi.org/10.1038/ng0197-30

Basson, C. T., Cowley, G. S., Solomon, S. D., Weissman, B., Poznanski, A. K., Traill, T. A., … Seidman, C. E. (1994). The Clinical and Genetic Spectrum of the Holt-Oram Syndrome (Heart-Hand Syndrome). *New England Journal of Medicine*, *330*(13), 885–891. https://doi.org/10.1056/NEJM199403313301302

Basson, C. T., Huang, T., Lin, R. C., Bachinsky, D. R., Weremowicz, S., Vaglio, A., … Seidman, C. E. (1999). Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(6), 2919–24. https://doi.org/10.1073/pnas.96.6.2919

Basson, C. T., Solomon, S. D., Weissman, B., Macrae, C. A., Andrew, K., Prieto, F., … Seidman, C. E. (1995). Genetic Heterogeneity of Heart-Hand Syndromes.

Baujat, G., & Le Merrer, M. (2007). Ellis-Van Creveld syndrome, *5*, 1–5. https://doi.org/10.1186/1750-1172-2-27

Becker, S. M., Al Halees, Z., Molina, C., & Paterson, R. M. (2001). Consanguinity and congenital heart disease in Saudi Arabia. *American Journal of Medical Genetics*, *99*(1), 8–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11170087

Bev Guo, M.D., Steven K. Lee, M.D., and Nader Paksima, D.O., M. P. H. (2013). A Review. *Bulletin of the Hospital for Joint Diseases*, *71*(1), 17–23.

Biesecker, L. G. (2011). Polydactyly: How many disorders and how many genes: 2010 update. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *240*(5), 931. https://doi.org/10.1002/dvdy.22609

Biesecker, L. G. (2012). Update, *240*(5), 931–942. https://doi.org/10.1002/dvdy.22609.Polydactyly

Bitar, F. F., Diab, K. A., Sabbagh, M., Siblini, G., & Obeid, M. (2001). Cardiac disease in children in Lebanon: The AUB-MC Children's Cardiac Registry Experience. *Journal Medical Libanais*, *49*(6), 304–310.

Bitar et al. Bitar et al 1999.pdf (1999).

Bittles, A. (2011). Assessing the influence of consanguinity on congenital heart disease. *Annals of Pediatric Cardiology*, *4*(2), 111–6. https://doi.org/10.4103/0974-2069.84637

Bizarro, R. O., & Callahan, J. A. (1970). Familial Atrial Septal Defect with Prolonged Atrioventricular Conduction Pattern of Inheritance. *Circulation*, 677–683.

Black, B. (2007). NIH Public Access. *Computer*, *144*(5), 724–732. https://doi.org/10.1038/jid.2014.371

Black, B. L., & Cripps, R. M. (2010). Myocyte Enhancer Factor 2 Transcription Factors in Heart Development and Disease. *Heart Development and Regeneration*, 673–699. https://doi.org/10.1016/B978-0-12-381332-9.00030-X

Blue, G. M., Kirk, E. P., Sholler, G. F., Harvey, R. P., & Winlaw, D. S. (2012). Congenital heart disease: Current knowledge about causes and inheritance. *Medical Journal of Australia*, *197*(3), 155–159. https://doi.org/10.5694/mja12.10811

Blue, G. M., & Winlaw, D. S. (2015). Next Generation Sequencing in Congenital Heart Disease : Gene Discovery and Clinical Application. *Next Generation: Sequencing & Applications*, *2*(1), 2–5. https://doi.org/10.4172/jngsa.1000113

Bogarapu, S., Bleyl, S. B., Calhoun, A., Viskochil, D., Saarel, E. V., Everitt, M. D., & Frank, D. U. (2014). Phenotype of a patient with contiguous deletion of TBX5 and TBX3: Expanding the disease spectrum. *American Journal of Medical Genetics, Part A*, *164*(5), 1304–1309. https://doi.org/10.1002/ajmg.a.36447

Bonnet, D., Pelet, A., Legeai-Mallet, L., Sidi, D., Mathieu, M., Parent, P., … Lyonnet, S. (1994). A gene for Holt?Oram syndrome maps to the distal long arm of chromosome 12. *Nature Genetics*, *6*(4), 405–408. https://doi.org/10.1038/ng0494-405

Brassington, A. M., Sung, S. S., Toydemir, R. M., Le, T., Roeder, A. D., Rutherford, A. E., … Bamshad, M. J. (2003). Expressivity of Holt-Oram syndrome is not predicted by TBX5 genotype. *American Journal of Human Genetics*, *73*(1), 74–85. https://doi.org/10.1086/376436

Breckpot, J., Thienpont, B., Arens, Y., Tranchevent, L. C., Vermeesch, J. R., Moreau, Y., … Devriendt, K. (2011). Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. *Cytogenetic and Genome Research*, *135*(3–4), 251–259. https://doi.org/10.1159/000331272

Breckpot, J., Thienpont, B., Peeters, H., de Ravel, T., Singer, A., Rayyan, M., … Devriendt, K. (2010). Array comparative genomic hybridization as a diagnostic tool for syndromic heart defects. *J Pediatr*, *156*(5), 810–7, 817 e1–817 e4. https://doi.org/10.1016/j.jpeds.2009.11.049

Brennan, J., Lu, C. C., Norris, D. P., Rodriguez, T. A., Beddington, R. S., & Robertson, E. J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. *Nature*, *411*(6840), 965–969. https://doi.org/10.1038/35082103

Bruneau, B. G. (2002). Transcriptional regulation of vertebrate cardiac morphogenesis. *Circulation Research*, *90*(5), 509–519. https://doi.org/10.1161/01.RES.0000013072.51957.B7

Bruneau, B. G. (2008). The developmental genetics of congenital heart disease. *Nature*, *451*(7181), 943–948. https://doi.org/10.1038/nature06801

Bruneau, B. G., Logan, M., Davis, N., Levi, T., Tabin, C. J., Seidman, J. G., & Seidman, C. E. (1999). Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. *Developmental Biology*, *211*(1), 100–8. https://doi.org/10.1006/dbio.1999.9298

Bruneau, B. G., Nemer, G., Schmitt, J. P., Charron, F., Robitaille, L., Caron, S., … Seidman, J. G. (2001). A murine model of Holt-Oram syndrome defines roles of the T-Box transcription factor Tbx5 in cardiogenesis and disease. *Cell*, *106*(6), 709–721. https://doi.org/10.1016/S0092-8674(01)00493-7

Cai, C.-L., Liang, X., Shi, Y., Chu, P.-H., Pfaff, S. L., Chen, J., & Evans, S. (2003). Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Developmental Cell*, *5*(6), 877–89. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14667410

Calvert, P. a, Rana, B. S., Kydd, A. C., & Shapiro, L. M. (2011). Patent foramen ovale: anatomy, outcomes, and closure. *Nature Reviews. Cardiology*, *8*(3), 148–160. https://doi.org/10.1038/nrcardio.2010.224

Camarata, T., Bimber, B., Kulisz, A., Chew, T. L., Yeung, J., & Simon, H. G. (2006). LMP4 regulates Tbx5 protein subcellular localization and activity. *Journal of Cell Biology*, *174*(3), 339–348. https://doi.org/10.1083/jcb.200511109

Camarata, T., Krcmery, J., Snyder, D., Park, S., Topczewski, J., & Simon, H. G. (2010). Pdlim7 (LMP4) regulation of Tbx5 specifies zebrafish heart atrio-ventricular boundary and valve formation. *Developmental Biology*, *337*(2), 233–245. https://doi.org/10.1016/J.YDBIO.2009.10.039

Camarata, T., Snyder, D., Schwend, T., Klosowiak, J., Holtrup, B., & Simon, H.-G. (2010). Pdlim7 is required for maintenance of the mesenchymal/epidermal Fgf signaling feedback loop during zebrafish pectoral fin development. *BMC Developmental Biology*, *10*(1), 104. https://doi.org/10.1186/1471-213X-10-104

Carmeliet, P., Ferreira, V., Breier, G., Pollefeyt, S., Kieckens, L., Gertsenstein, M., … Nagy, A. (1996). Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature*, *380*(6573), 435–439. https://doi.org/10.1038/380435a0

Cattaruzza, M., Lattrich, C., & Hecker, M. (2004). Focal Adhesion Protein Zyxin is a Mechanosensitive Modulator of Gene Expression in Vascular Smooth Muscle Cells. *Hypertension*, *43*(4), 726–730. https://doi.org/10.1161/01.HYP.0000119189.82659.52

Chaiyasap, P., Kulawonganunchai, S., Srichomthong, C., Tongsima, S., Suphapeetiporn, K., & Shotelersuk, V. (2014). Whole genome and exome sequencing of monozygotic twins with trisomy 21, discordant for a congenital heart defect and epilepsy. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0100191

Chang, D. F., Belaguli, N. S., Chang, J., & Schwartz, R. J. (2007). LIM-only protein, CRP2, switched on smooth muscle gene activity in adult cardiac myocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 157–62. https://doi.org/10.1073/pnas.0605635103

Chang, D. F., Belaguli, N. S., Iyer, D., Roberts, W. B., Wu, S.-P., Dong, X.-R., … Schwartz, R. J. (2003a). Cysteine-Rich LIM-Only Proteins CRP1 and CRP2 Are Potent Smooth Muscle Differentiation Cofactors. *Developmental Cell*, *4*(1), 107–118. https://doi.org/10.1016/S1534-5807(02)00396-9

Chang, D. F., Belaguli, N. S., Iyer, D., Roberts, W. B., Wu, S. P., Dong, X. R., … Schwartz, R. J. (2003b). Cysteine-rich LIM-only proteins CRP1 and CRP2 are potent smooth muscle differentiation cofactors. *Developmental Cell*, *4*(1), 107–118. https://doi.org/10.1016/S1534-5807(02)00396-9

Charron, F., Paradis, P., Bronchain, O., Nemer, G., & Nemer, M. (1999). Cooperative interaction between GATA-4 and GATA-6 regulates myocardial gene expression. *Molecular and Cellular Biology*, *19*(6), 4355–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10330176%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC104395

Chehab, G., Chedid, P., Saliba, Z., & Bouvagnet, P. (2007). Congenital cardiac disease and inbreeding: specific defects escape higher risk due to parental consanguinity. *Cardiology in the Young*, *17*(4), 414–22. https://doi.org/10.1017/S1047951107000704

Choi, M., Scholl, U. I., Ji, W., Liu, T., Tikhonova, I. R., Zumbo, P., … Lifton, R. P. (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(45), 19096–101. https://doi.org/10.1073/pnas.0910672106

Consortium, T. U. (2011). Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Research*, *39*(SUPPL. 1), 214–219. https://doi.org/10.1093/nar/gkq1020

Costain, G., Silversides, C. K., & Bassett, A. S. (2016). The importance of copy number variation in congenital heart disease. *Npj Genomic Medicine*, *1*(April), 16031. https://doi.org/10.1038/npjgenmed.2016.31

Coultas, L., Chawengsaksophak, K., & Rossant, J. (2005). Endothelial cells and VEGF in vascular development. *Nature*, *438*(7070), 937–945. https://doi.org/10.1038/nature04479

Crawford, A. W., Pino, J. D., & Beckerle, M. C. D. A.-J. (1994). Biochemical and Molecular Characterization of the Chicken Cysteine-Rich Protein, a Developmentally-Regulated Lim-Domain Protein That Is Associated With the Actin Cytoskeleton. *Journal of Cell Biology*, *124*(1–2), 117–127 ST–BIOCHEMICAL AND MOLECULAR CHARACTERI.

Dawid, I. B., Breen, J. J., & Toyama, R. (1998). LIM domains: Multiple roles as adapters and functional modifiers in protein interactions. *Trends in Genetics*, *14*(4), 156–162. https://doi.org/10.1016/S0168-9525(98)01424-3

Dice, J. E., & Bhatia, J. (2007). Patent ductus arteriosus: an overview. *The Journal of Pediatric Pharmacology and Therapeutics : JPPT : The Official Journal of PPAG*, *12*(3), 138–46. https://doi.org/10.5863/1551-6776-12.3.138

Digilio, M. C., Versacci, P., Lepri, F., Baban, A., Dallapiccola, B., & Marino, B. (2011). Atrioventricular canal defect and associated genetic disorders: new insights into polydactyly syndromes. *Cardiogenetics*, *1*(1). https://doi.org/10.4081/cardiogenetics.2011.e7

Donofrio, M. T., Moon-Grady, A. J., Hornberger, L. K., Copel, J. A., Sklansky, M. S., Abuhamad, A., … Rychik, J. (2014). Diagnosis and treatment of fetal cardiac disease: A scientific statement from the american heart association. *Circulation*, *129*(21), 2183–2242. https://doi.org/10.1161/01.cir.0000437597.44550.5d

Dube, J. Y., Chapdelaine, P., Trahan, P. L., Deperthes, D., Frenette, G., & Tremblay, R. R. (1998). Abundant cysteine-rich protein-1 is localized in the stromal

compartment of the human prostate. *Arch Androl*, *40*(2), 109–115. https://doi.org/10.3109/01485019808987933

Ecarnot-Laubriet, a, De Luca, K., Vandroux, D., Moisant, M., Bernard, C., Assem, M., … Teyssier, J. R. (2000). Downregulation and nuclear relocation of MLP during the progression of right ventricular hypertrophy induced by chronic pressure overload. *Journal of Molecular and Cellular Cardiology*, *32*(12), 2385–2395. https://doi.org/10.1006/jmcc.2000.1269

Egbe, A., Uppu, S., Lee, S., Ho, D., & Srivastava, S. (2014). Prevalence of associated extracardiac malformations in the congenital heart disease population. *Pediatric Cardiology*, 1239–1245. https://doi.org/10.1007/s00246-014-0922-6

Erdel, M., & Weiskirchen, R. (1998). Assignment 1 of CSRP1 encoding the LIM domain protein CRP1 , to human chromosome 1q32 by fluorescence in situ hybridization, *11*, 10–11.

Erdogan, F., Larsen, L. a, Zhang, L., Tümer, Z., Tommerup, N., Chen, W., … Ullmann, R. (2008). High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *Journal of Medical Genetics*, *45*(11), 704–9. https://doi.org/10.1136/jmg.2008.058776

Fahed, A. C., Gelb, B. D., Seidman, J. G., & Seidman, C. E. (2013). Genetics of congenital heart disease: The glass half empty. *Circulation Research*, *112*(4), 707–720. https://doi.org/10.1161/CIRCRESAHA.112.300853

Fauser, S., Munz, M., & Besch, D. (2003). ONLINE MUTATION REPORT Further support for digenic inheritance in Bardet-Biedl syndrome. *J Med Genet*, *40*. Retrieved from http://www.jmedgenet.com/cgi/content/full/40/8/e104

Feng, Q., Song, W., Lu, X., Hamilton, J. A., Lei, M., Peng, T., & Yee, S. P. (2002). Development of heart failure and congenital septal defects in mice lacking endothelial nitric oxide synthase. *Circulation*, *106*(7), 873–879. https://doi.org/10.1161/01.CIR.0000024114.82981.EA

Ferrara, N., Carver-Moore, K., Chen, H., Dowd, M., Lu, L., O'Shea, K. S., … Moore, M. W. (1996). Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature*, *380*(6573), 439–42. https://doi.org/10.1038/380439a0

Fish, E. W., Parnell, S. E., Sulik, K. K., Baker, L. K., Murdaugh, L. B., Lamson, D., & Williams, K. P. (2016). Preaxial polydactyly following early gestational exposure to the smoothened agonist, SAG, in C57BL/6J mice. *Birth Defects Research Part A - Clinical and Molecular Teratology*. https://doi.org/10.1002/bdra.23571

Flick, M. J., & Konieczny, S. F. (2000). The muscle regulatory and structural protein MLP is a cytoskeletal binding partner of betaI-spectrin. *Journal of Cell Science*, *113*, 1553–1564.

Förstermann, U., & Münzel, T. (2006). Endothelial nitric oxide synthase in vascular disease: From marvel to menace. *Circulation*, *113*(13), 1708–1714. https://doi.org/10.1161/CIRCULATIONAHA.105.602532

Freyd, G., Kim, S. K., & Horvitz, H. R. (1990). Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-

11. *Nature*. https://doi.org/10.1038/344876a0

Gao, X., Sedgwick, T., Shi, Y. B., & Evans, T. (1998). Distinct functions are implicated for the GATA-4, -5, and -6 transcription factors in the regulation of intestine epithelial cell differentiation. *Molecular and Cellular Biology*, *18*(5), 2901–2911.

Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., … Srivastava, D. (2003a). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature*, *424*(6947), 443–447. https://doi.org/10.1038/nature01827

Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., … Srivastava, D. (2003b). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature*, *424*(6947), 443–7. https://doi.org/10.1038/nature01827

Garg, V., Muth, A. N., Ransom, J. F., Schluterman, M. K., Barnes, R., King, I. N., … Srivastava, D. (2005). Mutations in NOTCH1 cause aortic valve disease. *Nature*, *437*(7056), 270–274. https://doi.org/10.1038/nature03940

Gelb, B. D., & Chung, W. K. (2014). Complex Genetics and the Etiology of Human Congenital Heart Disease. https://doi.org/10.1101/cshperspect.a013953

Geva, T., Martins, J. D., & Wald, R. M. (2014). Atrial septal defects. *The Lancet*, *383*(9932), 1921–1932. https://doi.org/10.1016/S0140-6736(13)62145-5

Ghosh, T. K., Packham, E. A., Bonser, A. J., Robinson, T. E., Cross, S. J., & Brook, J. D. (2001). Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome. *Human Molecular Genetics*, *10*(18), 1983–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11555635

Giuliano, P., Marino, R., Pinto, M. R., & De Santis, R. (1998). Identification and developmental expression of Ci-isl, a homologue of vertebrate islet genes, in the ascidian Ciona intestinalis. *Mechanisms of Development*, *78*(1–2), 199–202. https://doi.org/10.1016/S0925-4773(98)00143-9

Goldmuntz, E., Geiger, E., & Benson, D. W. (2001). NKX2.5 mutations in patients with tetralogy of fallot. *Circulation*, *104*(21), 2565–2568. https://doi.org/10.1161/hc4601.098427

Gorbach, D., Mote, B., Totir, L., Fernando, R., & Rothschild, M. (2010). Polydactyl inheritance in the pig. *Journal of Heredity*, *101*(4), 469–475. https://doi.org/10.1093/jhered/esq037

Grépin, C., Dagnino, L., Robitaille, L., Haberstroh, L., Antakly, T., & Nemer, M. (1994). A hormone-encoding gene identifies a pathway for cardiac but not skeletal muscle gene transcription. *Molecular and Cellular Biology*, *14*(5), 3115–3129. https://doi.org/10.1128/mcb.14.5.3115

Greulich, F., Rudat, C., & Kispert, A. (2011). Mechanisms of T-box gene function in the developing heart. *Cardiovascular Research*, *91*(2), 212–222. https://doi.org/10.1093/cvr/cvr112

Harvey, R. P. (2002). Patterning the vertebrate heart. *Nature Reviews. Genetics*, *3*(7), 544–556. https://doi.org/10.1038/nrg843

Hatcher, C. J., & Basson, C. T. (2001). Getting the T-box dose right. *Nature Medicine*, *7*(11), 1185–1186. https://doi.org/10.1038/nm1101-1185

Hatcher, C. J., & Basson, C. T. (2001). Getting the T-box dose right. *Nature Medicine*, *7*(11), 1185–6. https://doi.org/10.1038/nm1101-1185

He, A., Kong, S. W., Ma, Q., & Pu, W. T. (2011). Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(14), 5632–5637. https://doi.org/10.1073/pnas.1016959108

He, C., Cheng, H., & Zhou, R. (2007). GATA family of transcription factors of vertebrates: phylogenetics and chromosomal synteny. *Journal of Biosciences*, *32*(7), 1273–80. https://doi.org/10.1007/s12038-007-0127-8

Henderson, J. R., Brown, D., Richardson, J. a, Olson, E. N., & Beckerle, M. C. (2002). Expression of the gene encoding the LIM protein CRP2: a developmental profile. *The Journal of Histochemistry and Cytochemistry : Official Journal of the Histochemistry Society*, *50*(1), 107–11. https://doi.org/10.1177/002215540205000112

Henderson, J. R., Macalma, T., Brown, D., Richardson, J. A., Olson, E. N., & Beckerle, M. C. (1999). The LIM protein, CRP1, is a smooth muscle marker. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *214*(3), 229–38. https://doi.org/10.1002/(SICI)1097-0177(199903)214:3<229::AID-AJA6>3.0.CO;2-S

Hills, C. B., Kochilas, L., Schimmenti, L. A., & Moller, J. H. (2011). Ellis–van Creveld Syndrome and Congenital Heart Defects: Presentation of an Additional 32 Cases. *Pediatric Cardiology*, *32*(7), 977–982. https://doi.org/10.1007/s00246-011-0006-9

Hirayama-Yamada, K., Kamisago, M., Akimoto, K., Aotsuka, H., Nakamura, Y., Tomita, H., … Matsuoka, R. (2005). Phenotypes with GATA4 or NKX2.5 mutations in familial atrial septal defect. *American Journal of Medical Genetics*, *135 A*(1), 47–52. https://doi.org/10.1002/ajmg.a.30684

Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R., & Komuro, I. (2001). Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nature Genetics*, *28*(3), 276–80. https://doi.org/10.1038/90123

Hitz, M. P., Lemieux-Perreault, L. P., Marshall, C., Feroz-Zada, Y., Davies, R., Yang, S. W., … Andelfinger, G. (2012). Rare Copy Number Variants Contribute to Congenital Left-Sided Heart Disease. *PLoS Genetics*, *8*(9). https://doi.org/10.1371/journal.pgen.1002903

Hoffman, J. I. E., & Kaplan, S. (2002). The incidence of congenital heart disease. *Journal of the American College of Cardiology*, *39*(12), 1890–1900. https://doi.org/10.1016/S0735-1097(02)01886-7

Hoffman, J. I. E., Kaplan, S., & Liberthson, R. R. (2004). Prevalence of congenital heart disease. *American Heart Journal*, *147*(3), 425–439. https://doi.org/10.1016/j.ahj.2003.05.003

Horry, M. A., & Publishers, N. (2006). Progress In Medical Genetics Book Download.

HOUWELING, A., VANBORREN, M., MOORMAN, A., & CHRISTOFFELS, V.

(2005). Expression and regulation of the atrial natriuretic factor encoding gene during development and disease. *Cardiovascular Research*, *67*(4), 583–593. https://doi.org/10.1016/j.cardiores.2005.06.013

Huang, T. (2002). Current advances in Holt-Oram syndrome. *Curr Opin Pediatr*, *14*(6), 691–695. https://doi.org/10.1097/00008480-200212000-00008

Hummel, M., Li, Z. G., Pfaffinger, D., Neven, L., & Scanu, A. M. (1990). Familial hypercholesterolemia in a rhesus monkey pedigree: molecular basis of low density lipoprotein receptor deficiency. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(8), 3122–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2326270

Hwang, S., Park, B., Hwang, M., Kim, H., Kim, M., & Lee, J. (2016). Postaxial Polydactyly of the Hands and Feet in Monozygotic Twins, *21*(2), 48735.

Jakubek, Y., Lang, W., Vattathil, S., Garcia, M., Xu, L., Huang, L., … Kadara, H. (2016). Genomic landscape established by allelic imbalance in the cancerization field of a normal appearing airway. *Cancer Research*, *76*(13), 3676–3683. https://doi.org/10.1158/0008-5472.CAN-15-3064

Järvinen, P. (2011). *MOLECULAR STUDIES ON LIM-DOMAIN PROTEIN CRP1*.

Kadrmas, J. L., & Beckerle, M. C. (2004). The LIM domain: from the cytoskeleton to the nucleus. *Nature Reviews. Molecular Cell Biology*, *5*(11), 920–931. https://doi.org/10.1038/nrm1499

Kaiser, F. J., Brega, P., Raff, M. L., Byers, P. H., Gallati, S., Kay, T. T., … Lüdecke, H.-J. (2004). Novel missense mutations in the TRPS1 transcription factor define the nuclear localization signal. *European Journal of Human Genetics : EJHG*, *12*(2), 121–126. https://doi.org/10.1038/sj.ejhg.5201094

Kang, J. G., Sung, H. J., Jawed, S. I., Brenneman, C. L., Rao, Y. N., Sher, S., … Hwang, P. M. (2010). FOS expression in blood as a LDL-independent marker of statin treatment. *Atherosclerosis*, *212*(2), 567–570. https://doi.org/10.1016/j.atherosclerosis.2010.06.023

Khurana, T., Khurana, B., & Noegel, a a. (2002). LIM proteins: association with the actin cytoskeleton. *Protoplasma*, *219*(1–2), 1–12. https://doi.org/10.1007/s007090200000

Klaus, A., Saga, Y., Taketo, M. M., Tzahor, E., & Birchmeier, W. (2007). Distinct roles of Wnt/beta-catenin and Bmp signaling during early cardiogenesis. *Proceedings of the National Academy of Sciences*, *104*(47), 18531–18536. https://doi.org/10.1073/pnas.0703113104

Knöll, R., Hoshijima, M., Hoffman, H. M., Person, V., Lorenzen-Schmidt, I., Bang, M. L., … Chien, K. R. (2002). The cardiac mechanical stretch sensor machinery involves a Z disc complex that is defective in a subset of human dilated cardiomyopathy. *Cell*, *111*(7), 943–955. https://doi.org/10.1016/S0092-8674(02)01226-6

Kodo, K., Nishizawa, T., Furutani, M., Arai, S., Yamamura, E., Joo, K., … Yamagishi, H. (2009). GATA6 mutations cause human cardiac outflow tract defects by disrupting semaphorin-plexin signaling. *Proceedings of the National Academy of*

Sciences of the United States of America, *106*(33), 13933–8. https://doi.org/10.1073/pnas.0904744106

Kohlhase, J., Heinrich, M., Liebers, M., Fröhlich Archangelo, L., Reardon, W., & Kispert, A. (2002). Cloning and expression analysis of Sall4, the murine homologue of the gene mutated in Okihiro syndrome. *Cytogenetic and Genome Research*, *98*(4), 274–277. https://doi.org/10.1159/000071048

Kohlhase, J., Heinrich, M., Schubert, L., Liebers, M., Kispert, A., Laccone, F., … Reardon, W. (2002). Okihiro syndrome is caused by SALL4 mutations. *Human Molecular Genetics*, *11*(23), 2979–2987.

Kohlhase, J., Schubert, L., Liebers, M., Rauch, A., Becker, K., Mohammed, S. N., … Reardon, W. (2003). Mutations at the SALL4 locus on chromosome 20 result in a range of clinically overlapping phenotypes, including Okihiro syndrome, Holt-Oram syndrome, acro-renal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. *Journal of Medical Genetics*, *40*(7), 473–478. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12843316

Kong, Y., Flick, M. J., Kudla, a J., & Konieczny, S. F. (1997). Muscle LIM protein promotes myogenesis by enhancing the activity of MyoD. *Molecular and Cellular Biology*, *17*(8), 4750–4760.

Koshiba-takeuchi, K., Takeuchi, J. K., Arruda, E. P., Kathiriya, I. S., Mo, R., Hui, C., … Bruneau, B. G. (2006). Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart, *38*(2), 175–183. https://doi.org/10.1038/ng1707

Krause, A., Zacharias, W., Camarata, T., Linkhart, B., Law, E., Lischke, A., … Simon, H. G. (2004). Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Developmental Biology*, *273*(1), 106–120. https://doi.org/10.1016/j.ydbio.2004.05.024

Kunath, M., Lüdecke, H. J., & Vortkamo, A. (2002). Expression of Trps1 during mouse embryonic development. *Gene Expression Patterns*, *2*(1–2), 119–122. https://doi.org/10.1016/S0925-4773(02)00300-3

Kuroda, S., Tokunaga, C., Kiyohara, Y., Higuchi, O., Konishi, H., Mizuno, K., … Kikkawa, U. (1996). Protein-Protein Interaction of Zinc Finger Protein-Protein Interaction of Zinc Finger LIM Domains with Protein Kinase C *, 31029–31033. https://doi.org/10.1074/jbc.271.49.31029

Kwon, C., Arnold, J., Hsiao, E. C., Taketo, M. M., Conklin, B. R., & Srivastava, D. (2007). Canonical Wnt signaling is a positive regulator of mammalian cardiac progenitors. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(26), 10894–9. https://doi.org/10.1073/pnas.0704044104

Kwon, C., Qian, L., Cheng, P., Nigam, V., Arnold, J., & Srivastava, D. (2009). A regulatory pathway involving Notch1/beta-catenin/Isl1 determines cardiac progenitor cell fate. *Nature Cell Biology*, *11*(8), 951–7. https://doi.org/10.1038/ncb1906

Lage, K., Greenway, S. C., Rosenfeld, J. a, Wakimoto, H., Gorham, J. M., Segre, a V, … Larsen, L. a. (2012). Genetic and environmental risk factors in congenital heart disease functionally converge in protein networks driving heart development. *Proc Natl Acad Sci U S A*, *109*(35), 14035–14040. https://doi.org/10.1073/pnas.1210730109

Lalani, S. R., Shaw, C., Wang, X., Patel, A., Patterson, L. W., Kolodziejska, K., … Belmont, J. W. (2012). Rare DNA copy number variants in cardiovascular malformations with extracardiac abnormalities. *European Journal of Human Genetics*, *21*(2), 173–181. https://doi.org/10.1038/ejhg.2012.155

Lander, J., & Ware, S. M. (2014). Copy Number Variation in Congenital Heart Defects. *Current Genetic Medicine Reports*, *2*(3), 168–178. https://doi.org/10.1007/s40142-014-0049-3

Lentjes, M. H., Niessen, H. E., Akiyama, Y., de Bruïne, A. P., Melotte, V., & van Engeland, M. (2016). The emerging role of GATA transcription factors in development and disease. *Expert Reviews in Molecular Medicine*, *18*, e3. https://doi.org/10.1017/erm.2016.2

Lepore, J. J., Mericko, P. A., Cheng, L., Lu, M. M., Morrisey, E. E., & Parmacek, M. S. (2006). GATA-6 regulates semaphorin 3C and is required in cardiac neural crest for cardiovascular morphogenesis. *Journal of Clinical Investigation*, *116*(4), 929–939. https://doi.org/10.1172/JCI27363

Li, A., Ponten, F., & dos Remedios, C. G. (2012). The interactome of LIM domain proteins: The contributions of LIM domain proteins to heart failure and heart development. *Proteomics*, *12*(2), 203–225. https://doi.org/10.1002/pmic.201100492

Li, C., Fang, Z., Jiang, T., Zhang, Q., Liu, C., Zhang, C., & Xiang, Y. (2013). Serum microRNAs profile from genome-wide serves as a fingerprint for diagnosis of acute myocardial infarction and angina pectoris. *BMC Medical Genomics*, *6*, 16. https://doi.org/10.1186/1755-8794-6-16

Li, L., Krantz, I. D., Deng, Y., Genin, A., Banta, A. B., Collins, C. C., … Spinner, N. B. (1997). Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nat Genet*, *16*(3), 243–51. https://doi.org/10.1038/ng0797-243

Li, Q. Y., Newbury-Ecob, R. a, Terrett, J. a, Wilson, D. I., Curtis, a R., Yi, C. H., … Brook, J. D. (1997). Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nature Genetics*, *15*(1), 21–29. https://doi.org/10.1038/ng0197-21

Liberatore, C. M., Searcy-Schrick, R. D., Vincent, E. B., & Yutzey, K. E. (2002). Nkx-2.5 gene induction in mice is mediated by a Smad consensus regulatory region. *Developmental Biology*, *244*(2), 243–56. https://doi.org/10.1006/dbio.2002.0604

Liebhaber, S. A., Emery, J. G., Urbanek, M., Wang, X., & Cooke, N. E. (1990). Characterization of a human cDNA encoding a widely expressed and highly conserved cysteine-rich protein with an unusual zinc-finger motif. *Nucleic Acids Research*, *18*(13), 3871–3879. https://doi.org/10.1093/nar/18.13.3871

Lien, C. L., Wu, C., Mercer, B., Webb, R., Richardson, J. a, & Olson, E. N. (1999). Control of early cardiac-specific transcription of Nkx2-5 by a GATA-dependent enhancer. *Development (Cambridge, England)*, *126*, 75–84.

Lilly, B., Clark, K. A., Yoshigi, M., Pronovost, S., Wu, M. L., Periasamy, M., … Beckerle, M. C. (2010). Loss of the serum response factor cofactor, cysteine-rich protein 1, attenuates neointima formation in the mouse. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *30*(4), 694–701. https://doi.org/10.1161/ATVBAHA.109.200741

Lilly, B., Olson, E. N., & Beckerle, M. C. (2001). Identification of a CArG Box-Dependent Enhancer within the Cysteine-Rich Protein 1 Gene That Directs Expression in Arterial but Not Venous or Visceral Smooth Muscle Cells. *Developmental Biology*, *240*(2), 531–547. https://doi.org/10.1006/dbio.2001.0507

Lilly, B., Olson, E. N., & Beckerle, M. C. (2001). Identification of a CArG box-dependent enhancer within the cysteine-rich protein 1 gene that directs expression in arterial but not venous or visceral smooth muscle cells. *Developmental Biology*, *240*(2), 531–47. https://doi.org/10.1006/dbio.2001.0507

Lin, C.-J., Lin, C.-Y., Chen, C.-H., Zhou, B., & Chang, C.-P. (2012). Partitioning the heart: mechanisms of cardiac septation and valve development. *Development*, *139*(18), 3277–3299. https://doi.org/10.1242/dev.063495

Linden, H., Williams, R., King, J., Blair, E., & Kini, U. (2009). Ulnar mammary syndrome and *TBX3* : Expanding the phenotype. *American Journal of Medical Genetics Part A*, *149A*(12), 2809–2812. https://doi.org/10.1002/ajmg.a.33096

Linnemann, A. K., O'Geen, H., Keles, S., Farnham, P. J., & Bresnick, E. H. (2011). Genetic framework for GATA factor function in vascular biology. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13641–6. https://doi.org/10.1073/pnas.1108440108

Lipman, D. J., & Pearson, W. R. (1985). Rapid and sensitive protein similarity searches. *Science*, *227*(4693), 1435–1441. https://doi.org/2983426

Liu, W., & Foley, A. C. (2011). Signaling pathways in early cardiac development. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *3*(2), 191–205. https://doi.org/10.1002/wsbm.112

Loffredo, C. a, Chokkalingam, A., Sill, A. M., Boughman, J. a, Clark, E. B., Scheel, J., & Brenner, J. I. (2004). Prevalence of congenital cardiovascular malformations among relatives of infants with hypoplastic left heart, coarctation of the aorta, and d-transposition of the great arteries. *American Journal of Medical Genetics. Part A*, *124A*(3), 225–230. https://doi.org/10.1002/ajmg.a.20366

Louis, H. A., Pino, J. D., Schmeichel, K. L., Pomiès, P., & Beckerle, M. C. (1997). Comparison of three members of the cysteine-rich protein family reveals functional conservation and divergent patterns of gene expression. *Journal of Biological Chemistry*, *272*(43), 27484–27491. https://doi.org/10.1074/jbc.272.43.27484

Lyons, I., Parsons, L. M., Hartley, L., Li, R., Andrews, J. E., Robb, L., & Harvey, R. P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes and Development*, *9*(13), 1654–1666.

https://doi.org/10.1101/gad.9.13.1654

Maas, S. M., Shaw, A. C., Bikker, H., L??decke, H. J., van der Tuin, K., Badura-Stronka, M., … Hennekam, R. C. (2015a). Phenotype and genotype in 103 patients with tricho-rhino-phalangeal syndrome. *European Journal of Medical Genetics*, *58*(5), 279–292. https://doi.org/10.1016/j.ejmg.2015.03.002

Maas, S. M., Shaw, A. C., Bikker, H., L??decke, H. J., van der Tuin, K., Badura-Stronka, M., … Hennekam, R. C. (2015b). Phenotype and genotype in 103 patients with tricho-rhino-phalangeal syndrome. *European Journal of Medical Genetics*, *58*(5), 279–292. https://doi.org/10.1016/j.ejmg.2015.03.002

Maitra, M., Schluterman, M. K., Nichols, H. A., Richardson, J. A., Lo, C. W., Srivastava, D., & Garg, V. (2009). Interaction of Gata4 and Gata6 with Tbx5 is critical for normal cardiac development. *Developmental Biology*, *326*(2), 368–77. https://doi.org/10.1016/j.ydbio.2008.11.004

Malik, T. H., Shoichet, S. A., Latham, P., Kroll, T. G., Peters, L. L., & Shivdasani, R. A. (2001). Transcriptional repression and developmental functions of the atypical vertebrate GATA protein TRPS1. *EMBO Journal*, *20*(7), 1715–1725. https://doi.org/10.1093/emboj/20.7.1715

Mani, A., Meraji, S.-M., Houshyar, R., Radhakrishnan, J., Mani, A., Ahangar, M., … Lifton, R. P. (2002). Finding genetic contributions to sporadic disease: a recessive locus at 12q24 commonly contributes to patent ductus arteriosus. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(23), 15054–15059. https://doi.org/10.1073/pnas.192582999

Marian, A. J. (2017). Congenital Heart Disease. *Circulation Research*, *120*(6). Retrieved from http://circres.ahajournals.org/content/120/6/895.full

Marino, B., & Digilio, M. . (2000). Congenital Heart Disease and Genetic Syndromes. *Cardiovascular Pathology*, *9*(6), 303–315. https://doi.org/10.1016/S1054-8807(00)00050-8

Marino, B. S., Lipkin, P. H., Newburger, J. W., Peacock, G., Gerdes, M., Gaynor, J. W., … Mahle, W. T. (2012). Neurodevelopmental Outcomes in Children With Congenital Heart Disease: Evaluation and Management: A Scientific Statement From the American Heart Association. *Circulation*, *126*(9), 1143–1172. https://doi.org/10.1161/CIR.0b013e318265ee8a

Maturana, A. D., Nakagawa, N., Yoshimoto, N., Tatematsu, K., Hoshijima, M., Tanizawa, K., & Kuroda, S. (2011). LIM domains regulate protein kinase C activity: A novel molecular function. *Cellular Signalling*, *23*(5), 928–934. https://doi.org/10.1016/j.cellsig.2011.01.021

McCulley, D. J., & Black, B, L. (2012). Transcription Factor Pathways and Congenital Heart Disease. *Curr Top Dev Biol*, *100*(415), 253–277. https://doi.org/10.1016/B978-0-12-387786-4.00008-7.Transcription

McCulley, D. J., & Black, B. L. (2012). *Transcription Factor Pathways and Congenital Heart Disease*. *Current Topics in Developmental Biology* (Vol. 100). Elsevier Inc. https://doi.org/10.1016/B978-0-12-387786-4.00008-7

McDaniel, N. L. (2001). Ventricular and atrial septal defects. *Pediatrics in Review /*

American Academy of Pediatrics, *22*, 265–270. https://doi.org/10.1542/pir.22-8-265

Mcdermott, D. A., Bressan, M. C., He, J., Lee, J. S., Aftimos, S., Brueckner, M., … Basson, C. T. (2005). TBX5 genetic testing validates strict clinical criteria for Holt-Oram syndrome. *Pediatric Research*, *58*(5), 981–986. https://doi.org/10.1203/01.PDR.0000182593.95441.64

McKusick, V. a. (1992). Mendelian Inheritance in Man and its online version, OMIM. *American Journal of Human Genetics*, *80*(4), 588–604. https://doi.org/10.1086/514346

Meneghini, V., Odent, S., Platonova, N., Egeo, A., & Merlo, G. R. (2006). Novel TBX3 mutation data in families with Ulnar?Mammary syndrome indicate a genotype?phenotype relationship: mutations that do not disrupt the T-domain are associated with less severe limb defects. *European Journal of Medical Genetics*, *49*(2), 151–158. https://doi.org/10.1016/j.ejmg.2005.04.021

Miyasaka, K. Y., Kida, Y. S., Sato, T., Minami, M., & Ogura, T. (2007). Csrp1 regulates dynamic cell movements of the mesendoderm and cardiac mesoderm through interactions with Dishevelled and Diversin. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(27), 11274–11279. https://doi.org/10.1073/pnas.0702000104

Molkentin, J. D. (2000). The zinc finger-containing transcription factors GATA-4, -5, and -6: Ubiquitously expressed regulators of tissue-specific gene expression. *Journal of Biological Chemistry*, *275*(50), 38949–38952. https://doi.org/10.1074/jbc.R000029200

Molkentin, J. D., Kalvakolanu, D. V, & Markham, B. E. (1994). Transcription factor GATA-4 regulates cardiac muscle-specific expression of the alpha-myosin heavy-chain gene. *Mol.Cell Biol.*, *14*(7), 4947–4957. https://doi.org/10.1128/MCB.14.7.4947.Updated

Molkentin, J. D., Lin, Q., Duncan, S. A., & Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes and Development*, *11*(8), 1061–1072. https://doi.org/10.1101/gad.11.8.1061

Momeni, P., Glöckner, G., Schmidt, O., von Holtum, D., Albrecht, B., Gillessen-Kaesbach, G., … Lüdecke, H. J. (2000). Mutations in a new gene, encoding a zinc-finger protein, cause tricho-rhino-phalangeal syndrome type I. *Nature Genetics*, *24*(1), 71–74. https://doi.org/10.1038/71717

Mori, T., Chen, Y. F., Feng, J. A., Hayashi, T., Oparil, S., & Perry, G. J. (2004). Volume overload results in exaggerated cardiac hypertrophy in the atrial natriuretic peptide knockout mouse. *Cardiovascular Research*, *61*(4), 771–779. https://doi.org/10.1016/j.cardiores.2003.12.005

Morrisey, E. E., Ip, H. S., Tang, Z., & Parmacek, M. S. (1997). GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily. *Journal of Biological Chemistry*, *272*(13), 8515–8524. https://doi.org/10.1074/jbc.272.13.8515

Morrisey, E. E., Tang, Z., Sigrist, K., Lu, M. M., Jiang, F., Ip, H. S., & Parmacek, M. S.

(1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes and Development*, *12*(22), 3579–3590. https://doi.org/10.1101/gad.12.22.3579

Nabulsi, M. M., Tamim, H., Sabbagh, M., Obeid, M. Y., Yunis, K. a, & Bitar, F. F. (2003). Parental consanguinity and congenital heart malformations in a developing country. *American Journal of Medical Genetics. Part A*, *116A*(4), 342–7. https://doi.org/10.1002/ajmg.a.10020

Nadeau, M., Georges, R. O., Laforest, B., Yamak, A., Lefebvre, C., Beauregard, J., … Nemer, M. (2010). An endocardial pathway involving Tbx5, Gata4, and Nos3 required for atrial septum formation. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0914888107

Naito, A. T., Shiojima, I., Akazawa, H., Hidaka, K., Morisaki, T., Kikuchi, A., & Komuro, I. (2006). Developmental stage-specific biphasic roles of Wnt/beta-catenin signaling in cardiomyogenesis and hematopoiesis. *Proceedings of the National Academy of Sciences*, *103*(52), 19812–19817. https://doi.org/10.1073/pnas.0605768103

Najwer, I., & Lilly, B. (2005). Ca2+/calmodulin-dependent protein kinase IV activates cysteine-rich protein 1 through adjacent CRE and CArG elements. *American Journal of Physiology. Cell Physiology*, *289*(4), C785-93. https://doi.org/10.1152/ajpcell.00098.2005

Nemer, G., Fadlalah, F., Usta, J., Nemer, M., Dbaibo, G., Obeid, M., & Bitar, F. (2006). A novel mutation in theGATA4 gene in patients with Tetralogy of Fallot. *Human Mutation*, *27*(3), 293–294. https://doi.org/10.1002/humu.9410

Nemer, G., & Nemer, M. (2001). Regulation of heart development and function through combinatorial interactions of transcription factors. *Annals of Medicine*, *33*(9), 604–610. https://doi.org/10.3109/07853890109002106

Nemer, G., & Nemer, M. (2001). Regulation of heart development and function through combinatorial interactions of transcription factors. *Annals of Medicine*, *33*(9), 604–10. https://doi.org/10.3109/07853890109002106

Nemer, G., & Nemer, M. (2003). Transcriptional activation of BMP-4 and regulation of mammalian organogenesis by GATA-4 and -6. *Developmental Biology*, *254*(1), 131–148. https://doi.org/10.1016/S0012-1606(02)00026-X

Nemer, G., & Nemer, M. (2010). *GATA4 in Heart Development and Disease*. *Heart Development and Regeneration* (Vol. II). Elsevier Inc. https://doi.org/10.1016/B978-0-12-381332-9.00027-X

Nemer, G., Qureshi, S. T., Malo, D., & Nemer, M. (1999). Functional analysis and chromosomal mapping of Gata5, a gene encoding a zinc finger DNA-binding protein. *Mammalian Genome : Official Journal of the International Mammalian Genome Society*, *10*(10), 993–999. https://doi.org/10.1007/s003359901146

Nemer, M. (2008). Genetic insights into normal and abnormal heart development. *Cardiovascular Pathology*, *17*(1), 48–54. https://doi.org/10.1016/j.carpath.2007.06.005

Nemer G., B. F. (2006a). Progress in Medical Genetics - Marlene A. Horry - Google

Books. Retrieved June 23, 2017, from https://books.google.com.lb/books?id=fPIh_qPq_00C&pg=PA39&dq=Nemer+and+Bitar.+Molecular+Markers+of+Congenital+heart+disease&hl=en&sa=X&redir_esc=y#v=onepage&q=Nemer and Bitar. Molecular Markers of Congenital heart disease&f=false

Nemer G., B. F. (2006b). Progress in Medical Genetics - Marlene A. Horry - Google Books.

Newbury-Ecob, R., Leanage, R., Raebum, J. A., & Young, I. D. (1996). Holt-Oram syndrome :, 300–307.

Niederreither, K., Vermot, J., Messaddeq, N., Schuhbaur, B., Chambon, P., & Dollé, P. (2001). Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. *Development (Cambridge, England)*, *128*(7), 1019–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11245568

Nomir, A. G., Takeuchi, Y., Fujikawa, J., El Sharaby, A. A., Wakisaka, S., & Abe, M. (2016). Fate mapping of Trps1 daughter cells during cardiac development using novel Trps1-Cre mice. *Genesis*, *54*(7), 379–388. https://doi.org/10.1002/dvg.22951

Oka, T., Xu, J., & Molkentin, J. D. (2007). Re-employment of developmental transcription factors in adult heart disease. *Seminars in Cell and Developmental Biology*, *18*(1), 117–131. https://doi.org/10.1016/j.semcdb.2006.11.012

Olson, E. N. (2006). Gene regulatory networks in the evolution and development of the heart. *Science (New York, N.Y.)*, *313*(5795), 1922–7. https://doi.org/10.1126/science.1132292

Olson, E. N., & Nordheim, A. (2010). Linking actin dynamics and gene transcription to drive cellular motile functions. *Nature Reviews. Molecular Cell Biology*, *11*(5), 353–65. https://doi.org/10.1038/nrm2890

Packham, E. A., & Brook, J. D. (2003). T-box genes in human disorders. *Human Molecular Genetics*, *12*(90001), 37R–44. https://doi.org/10.1093/hmg/ddg077

Patent foramen ovale. (2015). *UpToDate*, 1–11. Retrieved from www.uptodate.com

Pathways, M. (2016). *Congenital Heart Diseases: The Broken Heart*. https://doi.org/10.1007/978-3-7091-1883-2

Peterkin, T., Gibson, A., Loose, M., & Patient, R. (2005). The roles of GATA-4, -5 and -6 in vertebrate heart development. *Seminars in Cell and Developmental Biology*, *16*(1), 83–94. https://doi.org/10.1016/j.semcdb.2004.10.003

Pfaff, S. L., Mendelsohn, M., Stewart, C. L., Edlund, T., & Jessell, T. M. (1996). Requirement for LIM homeobox gene Isl1 in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. *Cell*, *84*(2), 309–320. https://doi.org/10.1016/S0092-8674(00)80985-X

Pierpont, M. E., Basson, C. T., Benson, D. W., Gelb, B. D., Giglia, T. M., Goldmuntz, E., … Webb, C. L. (2007). Genetic Basis for Congenital Heart Defects: Current Knowledge A Scientific Statement From the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: Endorsed by the American Academy of Pediatrics. *Circulation*, *115*(23), 3015–3038. https://doi.org/10.1161/CIRCULATIONAHA.106.183056

Pikkarainen, S., Tokola, H., Kerkelä, R., & Ruskoaho, H. (2004). GATA transcription factors in the developing and adult heart. *Cardiovascular Research*, *63*(2), 196–207. https://doi.org/10.1016/j.cardiores.2004.03.025

Pomiès, P., Louis, H. A., & Beckerle, M. C. (1997). CRP1, a LIM domain protein implicated in muscle differentiation, interacts with ??-actinin. *Journal of Cell Biology*, *139*(1), 157–168. https://doi.org/10.1083/jcb.139.1.157

Postma, a V, Bezzina, C. R., & Christoffels, V. M. (2015). Genetics of congenital heart disease: the contribution of the noncoding regulatory genome. *J Hum Genet*, *61*(1), 13–19. https://doi.org/10.1038/jhg.2015.98

Prassler, J., Murr, a, Stocker, S., Faix, J., Murphy, J., & Marriott, G. (1998). DdLIM is a cytoskeleton-associated protein involved in the protrusion of lamellipodia in Dictyostelium. *Molecular Biology of the Cell*, *9*(3), 545–559.

Pu, W. T., Ishiwata, T., Juraszek, A. L., Ma, Q., & Izumo, S. (2004). GATA4 is a dosage-sensitive regulator of cardiac morphogenesis. *Developmental Biology*, *275*(1), 235–244. https://doi.org/10.1016/j.ydbio.2004.08.008

Ramegowda, S., & Ramachandra, N. B. (2005). Review Article An understanding the genetic basis of congenital heart disease, *11*(1), 14–23.

Ramegowda, S., & Ramachandra, N. B. (2006). Parental consanguinity increases congenital heart diseases in South India. *Annals of Human Biology*, *33*(5–6), 519–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17381050

Reiter, J. F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N., & Stainier, D. Y. R. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes and Development*, *13*(22), 2983–2995. https://doi.org/10.1101/gad.13.22.2983

Richards, A. A., & Garg, V. (2010). Genetics of congenital heart disease. *Current Cardiology Reviews*, *6*(2), 91–7. https://doi.org/10.2174/157340310791162703

Ruiz-Perez, V. L., Ide, S. E., Strom, T. M., Lorenz, B., Wilson, D., Woods, K., … Goodship, J. (2000). Mutations in a new gene in Ellis-van Creveld syndrome and Weyers acrodental dysostosis. *Nature Genetics*, *24*(3), 283–286. https://doi.org/10.1038/73508

Ruiz de la Fuente S, P. F. (1980). Heart-hand syndrome . III . A new syndrome in three, *55*(1), 43–47.

Sadler, I., Crawford, A. W., Michelsen, J. W., & Beckerle, M. C. (1992). Zyxin and cCRP: Two interactive LIM domain proteins associated with the cytoskeleton. *Journal of Cell Biology*, *119*(6), 1573–1587. https://doi.org/10.1083/jcb.119.6.1573

Sagave, J. F., Moser, M., Ehler, E., Weiskirchen, S., Stoll, D., Günther, K., … Weiskirchen, R. (2008). Targeted disruption of the mouse Csrp2 gene encoding the cysteine- and glycine-rich LIM domain protein CRP2 result in subtle alteration of cardiac ultrastructure. *BMC Developmental Biology*, *8*, 80. https://doi.org/10.1186/1471-213X-8-80

Sakai, Y., Nakagawa, R., Sato, R., & Maeda, M. (1998). Selection of DNA Binding Sites for Human Transcriptional Regulator GATA-6. *Biochemical and Biophysical*

Research Communications, *250*(3), 682–688. https://doi.org/10.1006/bbrc.1998.9374

Sánchez-García, I., Osada, H., Forster, A., & Rabbitts, T. H. (1993). The cysteine-rich LIM domains inhibit DNA binding by the associated homeodomain in Isl-1. *The EMBO Journal*, *12*(11), 4243–50. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=413719&tool=pmcentrez&rendertype=abstract

Schmeichel, K. L., & Beckerle, M. C. (1994). The LIM domain is a modular protein-binding interface. *Cell*, *79*(2), 211–219. https://doi.org/10.1016/0092-8674(94)90191-0

Schmeichel, K. L., & Beckerle, M. C. (1998). LIM domains of cysteine-rich protein 1 (CRP1) are essential for its zyxin-binding function. *The Biochemical Journal*, *331 (Pt 3*, 885–92. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1219431&tool=pmcentrez&rendertype=abstract

Searcy, R. D., Vincent, E. B., Liberatore, C. M., & Yutzey, K. E. (1998). A GATA-dependent nkx-2.5 regulatory element activates early cardiac gene expression in transgenic mice. *Development (Cambridge, England)*, *125*(22), 4461–4470. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9778505

Seidman, J. G., & Seidman, C. (2002). Transcription factor haploinsufficiency: When half a loaf is not enough. *Journal of Clinical Investigation*, *109*(4), 451–455. https://doi.org/10.1172/JCI200215043

Sepulveda, J. L., Vlahopoulos, S., Iyer, D., Belaguli, N., & Schwartz, R. J. (2002). Combinatorial expression of GATA4, Nkx2-5, and serum response factor directs early cardiac gene activity. *Journal of Biological Chemistry*, *277*(28), 25775–25782. https://doi.org/10.1074/jbc.M203122200

Shiaulou Yuan, Samir Zaidi, and M. B. (2014). Congenital Heart Disease: emerging themes linking genetics and development, *23*(3), 352–359. https://doi.org/10.1016/j.gde.2013.05.004.Congenital

Shieh, J. T. C., Bittles, A. H., & Hudgins, L. (2012). Consanguinity and the Risk of Congenital Heart Disease, (April), 1236–1241. https://doi.org/10.1002/ajmg.a.35272

Shott, J., Benson, D., Basson, C., Pease, W., Silberbach, G., Moak, J., & Maron, B. (1998). Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science*, *281*(5373), 108–111. https://doi.org/10.1126/science.281.5373.108

Silengo, M. C., Biagioli, M., Guala, A., Lopez-Bell, G., & Lala, R. (1990). Heart-hand syndrome II. A report of Tabatznik syndrome with new findings. *Clin Genet*, *38*(2), 105–113. https://doi.org/10.1111/j.1399-0004.1990.tb03556.x

Silversides, C. K., Kiess, M., Beauchesne, L., Bradley, T., Connelly, M., Niwa, K., … Therrien, J. (2010). Canadian Cardiovascular Society 2009 Consensus Conference on the management of adults with congenital heart disease: outflow tract obstruction, coarctation of the aorta, tetralogy of Fallot, Ebstein anomaly and

Marfan's syndrome. *The Canadian Journal of Cardiology*, *26*(3), e80-97. https://doi.org/10.1016/S0828-282X(10)70355-X

Silversides, C. K., Lionel, A. C., Costain, G., Merico, D., Migita, O., Liu, B., … Bassett, A. S. (2012). Rare Copy Number Variations in Adults with Tetralogy of Fallot Implicate Novel Risk Gene Pathways. *PLoS Genetics*, *8*(8). https://doi.org/10.1371/journal.pgen.1002843

Silversides, C. K., Salehian, O., Editor, S., Oechslin, E., Schwerzmann, M., Muhll, I. V., … Oechslin, E. (2010). SpeCial artiCle Canadian Cardiovascular Society 2009 Consensus Conference on the management of adults with congenital heart disease : Complex congenital cardiac lesions, *26*(3), 98–117.

Simon, M. C. (1995). Gotta have GATA. *Nature Genetics*, *11*(1), 9–11. https://doi.org/10.1038/ng0995-9

Sirbu, I. O., Zhao, X., & Duester, G. (2008). Retinoic acid controls heart anteroposterior patterning by down-regulating Isl1 through the Fgf8 pathway. *Developmental Dynamics*, *237*(6), 1627–1635. https://doi.org/10.1002/dvdy.21570

Slavotinek, A. M. (2017). McKusick-Kaufman Syndrome. *GeneReviews®*.

Slavotinek, A. M., & Biesecker, L. G. (2000). Phenotypic overlap of McKusick-Kaufman syndrome with Bardet-Biedl syndrome: A literature review. *American Journal of Medical Genetics*, *95*(3), 208–215. https://doi.org/10.1002/1096-8628(20001127)95:3<208::AID-AJMG5>3.0.CO;2-J

Small, E. M., & Krieg, P. A. (2003). Transgenic analysis of the atrialnatriuretic factor (ANF) promoter: Nkx2-5 and GATA-4 binding sites are required for atrial specific expression of ANF. *Developmental Biology*, *261*(1), 116–131. https://doi.org/10.1016/S0012-1606(03)00306-3

Snider, P., Fix, J. L., Rogers, R., Peabody-Dowling, G., Ingram, D., Lilly, B., & Conway, S. J. (2008). Generation and characterization of Csrp1 enhancer-driven tissue-restricted Cre-recombinase mice. *Genesis*, *46*(3), 167–176. https://doi.org/10.1002/dvg.20379

Soemedi, R., Wilson, I. J., Bentham, J., Darlay, R., Töpf, A., Zelenika, D., … Keavney, B. D. (2012). Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *American Journal of Human Genetics*, *91*(3), 489–501. https://doi.org/10.1016/j.ajhg.2012.08.003

Srivastava, D. (2001). G ENETIC A SSEMBLY OF THE H EART : Implications for Congenital Heart Disease.

Stronach, B. E., Siegrist, S. E., & Beckerle, M. C. (1996). Two muscle-specific LIM proteins in Drosophila. *Journal of Cell Biology*, *134*(5), 1179–1195. https://doi.org/10.1083/jcb.134.5.1179

Svensson, E. C., Huggins, G. S., Lin, H., Clendenin, C., Jiang, F., Tufts, R., … Leiden, J. M. (2000). A syndrome of tricuspid atresia in mice with a targeted mutation of the gene encoding Fog-2. *Nature Genetics*, *25*(3), 353–6. https://doi.org/10.1038/77146

Tamura, S., Wang, X. H., Maeda, M., & Futai, M. (1993). Gastric DNA-binding proteins recognize upstream sequence motifs of parietal cell-specific genes.

Proceedings of the National Academy of Sciences of the United States of America, *90*(22), 10876–10880. https://doi.org/10.1073/pnas.91.10.4609

Tartaglia, M., Kalidas, K., Shaw, A., Song, X., Musat, D. L., van der Burgt, I., … Gelb, B. D. (2002). PTPN11 Mutations in Noonan Syndrome: Molecular Spectrum, Genotype-Phenotype Correlation, and Phenotypic Heterogeneity. *The American Journal of Human Genetics*, *70*(6), 1555–1563. https://doi.org/10.1086/340847

Temtamy, S. A., & McKusick, V. (1978). The Genetics of Hand Malformations. *Birth Defects Original Article Series*, *14*(3), 332–333.

Tian, Y., Yuan, L., Goss, A. M., Wang, T., Yang, J., Lepore, J. J., … Edward, E. (2011). NIH Public Access, *18*(2), 275–287. https://doi.org/10.1016/j.devcel.2010.01.008.Characterization

Van Karnebeek, C. D. M., & Hennekam, R. C. M. (1999). Associations between chromosomal anomalies and congenital heart defects: A database search. *American Journal of Medical Genetics*, *84*(2), 158–166. https://doi.org/10.1002/(SICI)1096-8628(19990521)84:2<158::AID-AJMG13>3.0.CO;2-5

Ventricular septal defect. (2009). *The Surgical Clinics of North America*, *41*, 407–416.

Verheij, J. B. G. M., de Munnik, S. A., Dijkhuizen, T., de Leeuw, N., Olde Weghuis, D., van den Hoek, G. J., … van Ravenswaaij-Arts, C. M. A. (2009a). An 8.35??Mb overlapping interstitial deletion of 8q24 in two patients with coloboma, congenital heart defect, limb abnormalities, psychomotor retardation and convulsions. *European Journal of Medical Genetics*, *52*(5), 353–357. https://doi.org/10.1016/j.ejmg.2009.05.006

Verheij, J. B. G. M., de Munnik, S. A., Dijkhuizen, T., de Leeuw, N., Olde Weghuis, D., van den Hoek, G. J., … van Ravenswaaij-Arts, C. M. A. (2009b). An 8.35??Mb overlapping interstitial deletion of 8q24 in two patients with coloboma, congenital heart defect, limb abnormalities, psychomotor retardation and convulsions. *European Journal of Medical Genetics*, *52*(5), 353–357. https://doi.org/10.1016/j.ejmg.2009.05.006

Wang, D., Passier, R., Liu, Z.P., Shin, C.H., Wang, Z., Li, S., Sutherland, L.B., Small, E., Krieg, P.A. and Olson, E. N. (2002). Regulation of cardiac growth and development by SRF and its cofactors. - PubMed - NCBI. Retrieved June 15, 2017, from https://www.ncbi.nlm.nih.gov/pubmed/12858529

Wang, X., Leesv, G., Liebhaberslll, S. A., & Cookes, N. E. (1992). Human Cysteine-rich Protein, *267*(13).

Wang, X., Li, P., Chen, S., Xi, L., Guo, Y., Guo, A., & Sun, K. (2014). *Molecular medicine reports. Molecular Medicine Reports* (Vol. 9). D.A. Spandidos. Retrieved from https://www.spandidos-publications.com/10.3892/mmr.2013.1847

Warburton, D., Ronemus, M., Kline, J., Jobanputra, V., Williams, I., Anyane-Yeboa, K., … Levy, D. (2014). The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Human Genetics*, *133*(1), 11–27. https://doi.org/10.1007/s00439-013-1353-9

Weiskirchen, R., & Günther, K. (2003). The CRP/MLP/TLP family of LIM domain

proteins: Acting by connecting. *BioEssays*, *25*(2), 152–162. https://doi.org/10.1002/bies.10226

Weiskirchen, R., Pino, J. D., Macalma, T., Bister, K., & Beckerle, M. C. (1995). The cysteine-rich protein family of highly related LIM domain proteins. *Journal of Biological Chemistry*, *270*(48), 28946–28954. https://doi.org/10.1074/jbc.270.48.28946

Wilson, G. N. (1998). Correlated heart/limb anomalies in Mendelian syndromes provide evidence for a cardiomelic developmental field. *American Journal of Medical Genetics*, *76*(4), 297–305. https://doi.org/10.1002/(SICI)1096-8628(19980401)76:4<297::AID-AJMG4>3.0.CO;2-N

Xin, M., Davis, C. A., Molkentin, J. D., Lien, C.-L., Duncan, S. A., Richardson, J. A., & Olson, E. N. (2006). A threshold of GATA4 and GATA6 expression is required for cardiovascular development. *Proceedings of the National Academy of Sciences*, *103*(30), 11189–11194. https://doi.org/10.1073/pnas.0604604103

Xue, D., Tu, Y., & Chalfie, M. (1993). Cooperative interactions between the Caenorhabditis elegans homeoproteins UNC-86 and MEC-3. *Science*, *261*(5126), 1324–1328. https://doi.org/10.1126/science.8103239

Yang, L. (2011). From fibroblast cells to cardiomyocytes: direct lineage reprogramming. *Stem Cell Research & Therapy*, *2*(1), 1. https://doi.org/10.1186/scrt42

Yang, Y. Q., Li, L., Wang, J., Liu, X. Y., Chen, X. Z., Zhang, W., … Fang, W. Y. (2012). A novel GATA4 loss-of-function mutation associated with congenital ventricular septal defect. *Pediatric Cardiology*, *33*(4), 539–546. https://doi.org/10.1007/s00246-011-0146-y

Yang, Y. Q., Li, L., Wang, J., Liu, X. Y. X., Chen, X. Z., Zhang, W., … Bhuiyan, Z. A. (2012). High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *Circulation*, *45*(4), 704–9. https://doi.org/10.1136/jmg.2008.058776

Yao, X., Pérez-Alvarado, G. C., Louis, H. A., Pomiès, P., Hatt, C., Summers, M. F., & Beckerle, M. C. (1999). Solution Structure of the Chicken Cysteine-Rich Protein, CRP1, a Double-LIM Protein Implicated in Muscle Differentiation †,‡. *Biochemistry*, *38*(18), 5701–5713. https://doi.org/10.1021/bi982036y

Yet, S. F., Folta, S. C., Jain, M. K., Hsieh, C. M., Maemura, K., Layne, M. D., … Lee, M. E. (1998). Molecular cloning, characterization, and promoter analysis of the mouse Crp2/SmLim gene. Preferential expression of its promoter in the vascular smooth muscle cells of transgenic mice. *J Biol Chem*, *273*(17), 10530–10537.

Yin, Z., Haynie, J., Yang, X., Han, B., Kiatchoosakun, S., Restivo, J., … Yang, Y.-C. (2002). The essential role of Cited2, a negative regulator for HIF-1alpha, in heart development and neurulation. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(16), 10488–93. https://doi.org/10.1073/pnas.162371799

Yuasa, S., Itabashi, Y., Koshimizu, U., Tanaka, T., Sugimura, K., Kinoshita, M., … Fukuda, K. (2005). Transient inhibition of BMP signaling by Noggin induces

cardiomyocyte differentiation of mouse embryonic stem cells. *Nature Biotechnology*, *23*(5), 607–11. https://doi.org/10.1038/nbt1093

Yuval Dor1, Todd D. Camenisch2, Ahuva Itin1, Glenn I. Fishman3, John A. McDonald2, Peter Carmeliet4 and Eli Keshet1, *. (2001). VEGF and developmental heart defects. *Development*, *128*, 1531–1538. Retrieved from http://dev.biologists.org/content/develop/128/9/1531.full.pdf

Zaffran, S., Robrini, N. El, & Bertrand, N. (2014). Retinoids and Cardiac Development. *Journal of Developmental Biology*, *2*(Figure 1), 50–71. https://doi.org/10.3390/jdb2010050

Zeisberg, E. M., Ma, Q., Juraszek, A. L., Moses, K., Schwartz, R. J., Izumo, S., & Pu, W. T. (2005). Morphogenesis of the right ventricle requires myocardial expression of Gata4. *Journal of Clinical Investigation*, *115*(6), 1522–1531. https://doi.org/10.1172/JCI23769

Zhang, R., Thiele, H., Bartmann, P., Hilger, A. C., Berg, C., Herberg, U., … Reutter, H. (2016). Whole-Exome Sequencing in Nine Monozygotic Discordant Twins. *Twin Research and Human Genetics*, *19*(1), 60–65. https://doi.org/10.1017/thg.2015.93

Zheng, Q., & Zhao, Y. (2007). The diverse biofunctions of LIM domain proteins: determined by subcellular localization and protein-protein interaction. *Biology of the Cell / under the Auspices of the European Cell Biology Organization*, *99*(9), 489–502. https://doi.org/10.1042/BC20060126

Zhou, L., Liu, J., Xiang, M., Olson, P., Guzzetta, A., Zhang, K., … Xie, L. (2017). Gata4 potentiates second heart field proliferation and Hedgehog signaling for cardiac septation. *Proceedings of the National Academy of Sciences*, 201605137. https://doi.org/10.1073/PNAS.1605137114