MOLECULAR MECHANISMS OF *TBX20* REGULATION

Elizabeth Marie Mandel

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology (Molecular, Cellular and Developmental Biology).

Chapel Hill
2008

Approved By:

Dr. Frank L. Conlon

Dr. Victoria L. Bautch

Dr. William F. Marzluff

Dr. Sharon L. Milgram

Dr. W. Cam Patterson

ABSTRACT

ELIZABETH MARIE MANDEL: Molecular Mechanisms of *Tbx20* Regulation

(Under the direction of Dr. Frank L. Conlon)

The proper formation of the vertebrate heart involves the orchestration of many complex processes including induction, commitment, cell migration and morphogenesis. Each of these events is controlled by regulated gene expression within multiple transcriptional regulatory networks in a spatially and temporally restricted manner, and only slight changes in gene expression can lead to severe cardiac anomalies. One key example is provided by the T-box transcription factor, TBX20. Homologs of *Tbx20* have been described in a variety of organisms including mouse, zebrafish, chick, *Xenopus* and human, and with *Nkx2*.5 and *Tbx5*, it is one of the earliest markers of cardiac tissue. Analyses of TBX20 knockdown in mouse and *Xenopus* have also demonstrated a requirement for TBX20 in early heart development through an involvement in morphogenesis, cell proliferation and transcriptional regulation. However, the mechanisms controlling the tissue-specific regulation of *Tbx20* have yet to be described. Recently, both missense and nonsense mutations within the DNA-binding domain of human TBX20 have been shown to result in diverse cardiac pathologies including septal defects, valve defects, and cardiomyopathy. In an effort to clearly define the origins of such congenital heart defects, it will therefore be necessary to fully understand the transcriptional regulation and localized expression of genes such as *Tbx20*, that are involved in the earliest stages of cardiac development.

To this end, the work described here focuses on the characterization of *Tbx20* transcriptional regulation in the developing heart. We first demonstrate that the sequence and expression patterns of multiple T-box genes, including *Tbx20*, are conserved across species, including in the diploid frog *Xenopus tropicalis* (*X. tropicalis*). Our studies next go on to show that *Tbx20*, unlike other cardiac genes described to date, is expressed throughout all regions of the developing heart including both myocardial and endocardial layers. Utilizing transgenesis in both *Xenopus* and zebrafish, we have identified and characterized a conserved 334bp regulatory element that is sufficient to drive expression of *Tbx20* throughout these regions of the heart. To compliment these studies, we have also demonstrated a role for the BMP signaling pathway in the regulation of *Tbx20*, whereby the downstream BMP mediator SMAD1 directly binds regions within the cardiac regulatory element to drive gene expression in a dose-dependent fashion. This is the first evidence for a role of specific signaling pathways in the regulation of *Tbx20* expression, and provides key insights into the transcriptional regulatory network that is responsible for proper cardiac development.

ACKNOWLEDGEMENTS

"At times our own light goes out and is
rekindled by a spark from another person.
Each of us has cause to think with deep
gratitude of those who have lighted the flame within us."
~Albert Schweitzer

I am so very grateful to the many people who have helped to shape who I am, and who have guided me through my endeavors. First, I would like to thank my advisor, Frank Conlon. Frank has been supportive, motivating and a great role model to me over the past five years. He has taught me a great deal about research and developmental biology, and has pushed me to become a better scientist. I could ask for nothing more in a mentor and feel that his guidance during my graduate career has helped me to reach my goals.

I would also like to thank the members of the Conlon Lab, both past and present. I have enjoyed coming to work everyday not only because of the science, but also because I have been fortunate enough to work in such a great environment and with such great people. Sarah, Daniel, Chris and Jackie have been great role models and have given me much needed guidance as I follow in their footsteps. Yvette, Kathleen, Erika, Stephen and Shauna have spent countless hours listening to practice talks, offering advice, and making me smile on a daily basis. I have also had the wonderful opportunity to work with a number of talented undergraduates, Bahay, Brittney and Elizabeth. They have taught me a great deal, have helped me to become a better mentor and have made sure I've taken time each day to laugh.

My dissertation committee, Vicki Bautch, Bill Marzluff, Sharon Milgram and Cam Patterson, has been the source of much scientific guidance over the past 5 years. They have

offered many suggestions that have improved the quality of my work, and have helped me to think about my project from a variety of different perspectives.

My graduate school experience has been truly enriched by the friends I have been fortunate enough to meet in Chapel Hill, and by the friends who have continued to support me from across the country. Emily, Erin, Jade, Jenny, Bonnie, Pam, Lauren, Kelsea, the Criglers and the "Spaceman Spiffs" (and all of the others I don't have room to mention) have each found a unique way to make the past five years full of fun and good times.

I don't even know how to begin to properly thank my family for their love and support in everything that I do. I know that I can always count on my parents for advice, a listening ear, and unwavering love. They have always been there for me, regardless of the time or need. They are sources of inspiration, sources of encouragement, wonderful role models, and have made me who I am today. I am also grateful to have two terrific brothers, and a great sister-in-law. Craig, Tim and Mary are three of my best friends and have always been supportive, entertaining, understanding and grounding.

Lastly, I'd like to thank Cory. While we spent four years together at Knox, it wasn't until graduate school that we became close friends. It has been a great blessing to have someone with whom to share the highs and lows over the past five years. He has been supportive, understanding and my best friend.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

ABBREVIATIONS

ASD	Atrial Septal Defect
BAC	Bacterial Artificial Chromosome
BLAST	Basic Local Alignment and Search Tool
BMP	Bone Morphogenetic Protein
BNP	Brain Natriuretic Peptide
CD	Conduction System Defect
CHORI	Children's Hospital Oakland Research Institute
CSD	Cardiac Septal Defect
CHD	Congenital Heart Defect
CNS	Central Nervous System
Cx40	Connexin 40
DAPI	4',6-diamidino-2-phenylindole
DCM	Dilated Cardiomyopathy
CPVT	Catecholaminergic Polymorphic Ventricular Tachycardia
DM	Differentiation Medium
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
DSHB	Developmental Studies Hybridoma Bank
eHAND	Heart- and Neural Crest Derivatives-Expressed 1
EGF	Epidermal Growth Factor
EGFP	Enhanced Green Fluorescent Protein
ES	Embryonic Stem

EYA	Eyes Absent
5'-FAM	5' Carboxyfluorescein
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
FIGE	Field Inversion Gel Electrophoresis
FOG2	Friend of GATA
FOXc1/2	Human Forkhead-box Subfamily c1/2
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
GATA	GATA Binding Protien
GFP	Green Fluorescent Protein
GM	Growth Medium
GST	Glutathione *S*-Transferase
HAT	Histone Acetyltransferase
hCG	Human Chorionic Gonadotropin
HDAC	Histone Deacetylase
HEY2	Hairy/Enhancer of Split-Related with YRPW Motif-2
HLH	Hypoplastic Left Heart
HOS	Holt-Oram Syndrome
HPF	Hours Post Fertilization
IFT	Inflow Tract
IGF	Insulin-like Growth Factor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRX4	Iroquois 4

MEF2c	Myocyte Enhancer Factor 2c Isoform
α/β-MHC	Alpha/Beta-Myosin Heavy Chain
miRNA	Micro Ribonucleic Acid
MLC2V	Myosin Light Chain-2 Ventricular Isoform
MMLV	Maloney Murine Leukemia Virus
MO	Morpholino Oligonucleotide
MSX2	Muscle Segment Homeobox 2
N-CAM	Neural Cell Adhesion Molecule
NFATc4	Nuclear Factor of Activated T Cells, Cytoplasmic, Calcineurin-Dependent-4
NKX2.5	NK2-Related Homeobox
N-MYC	Neuroblastoma-Myelocytomatosis Viral-Related Oncogene
NPPA/ANF	Natriuretic Peptide Precursor A/Atrial Natriuretic Factor
OFT	Outflow Tract
PA	Pulmonary Atresia
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
REMI	Restriction Enzyme Mediated Integration
RLM-RACE	RNA Ligase-Mediated Rapid Amplification of cDNA Ends
RNA	Ribonucleic Acid
RT	Room Temperature
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SALL4	Spalt-Like 4

SBE	SMAD Binding Element
SHF	Secondary Heart Field
SMAD1	Mothers Against Decapentaplegic homolog 1
SNHL	Sensorineural Hearing Loss
SRF	Serum Response Factor
TA	Tricuspid Atresia, Truncus Arteriosus
TBX	T-box
TF	Transcription Factor
TFAP2B	Transcription Factor AP-2 beta
TGF	Transforming Growth Factor
TnC	Troponin C
TnI	Troponin I
TRA	Truncus Arteriosus
TOF	Tetralogy of Fallot
UTR	Untranslated Region
VEGF	Vascular Endothelial Growth Factor
VSD	Ventricular Septal Defects
XBRA	*Xenopus* Brachyury

CHAPTER 1

INTRODUCTION

Congenital heart defects (CHDs) threaten nearly 1% of all newborns and pose as a significant threat of infant death; however, the underlying genetic mechanisms of many CHDs remain elusive. Most likely, the majority of these defects have a basis in the complex process of cardiogenesis. Heart development is a progression of highly coordinated events including cell proliferation, differentiation, migration and morphogenesis, and it is likely that the alteration in expression of genes involved in these processes underlies a variety of cardiovascular defects. Therefore, by better understanding the molecular and cellular processes involved in normal heart development, we will gain insight into the origin of CHDs. To this end, a number of genes have been identified as potential causes of specific cardiac anomalies. For example, T-box genes, such as *Tbx1*, *Tbx5* are associated with congenital heart defects in humans in affiliation with DiGeorge syndrome and Holt-Oram syndrome phenotypes, respectively. Additionally, *Tbx20*, another member of the T-box gene family, has recently been linked to human congenital heart disease as nonsense or missense mutations in the genomic locus result in a complex spectrum of cardiovascular defects. In an effort to further understand its function and significance, *Tbx20* has been cloned and its expression characterized in numerous models, including *Xenopus*. Along with *Nkx2.5* and *Tbx5*, *Tbx20* is one of the earliest markers of heart tissue in *Xenopus* and appears

to play important roles in early heart development in a variety of organisms. However, the regulation and transcriptional regulatory functions of *Tbx20* have yet to be described.

Xenopus as a model system for studies of early heart development

Supported by over one hundred years of experimental embryology in amphibians, *Xenopus* is a well-established and outstanding model in which to study early organogenesis (67). *Xenopus* embryos are large, easily manipulated and easily obtained in large numbers. In addition, a well established fate map of embryos beginning at the gastrula stage allows for the identification of cells which will give rise to the heart as early as the 32-cell stage (16, 37). As *Xenopus* embryos are non-placental, and raised in liquid media, early development can proceed even in the absence of a functional circulatory system, allowing for a more detailed analysis of the mechanisms of early cardiovascular development. Although there are significant differences between the anatomy of adult amphibian and human hearts, it appears that many of the molecular and cellular processes involved in early heart development are conserved (5, 13, 67).

In addition, the use of the diploid frog, *Xenopus tropicalis* (*X. tropicalis*), has recently made genetic analysis feasible. These frogs reach sexual maturity in as little as 6 months, and can be utilized in both embryological and genetic studies. Additionally, the *X. tropicalis* genome has been fully sequenced, partially annotated and is publicly available (Joint Genome Institute, U.S. Department of Energy). In combination with the well established *Xenopus laevis* (*X. laevis*) system, the *X. tropicalis* system will be a very important tool in studies of the specification of cardiac precursor cells and early development (55, 67).

Early Vertebrate Heart Development

Morphogenesis

In vertebrates, the heart represents the first functioning organ. During early embryogenesis, heart precursors originate from bilateral patches of cells in the anterior lateral plate mesoderm that include populations of both myocardial and endocardial cells (15, 62). These cells are specified to the cardiac lineage during the process of gastrulation (3, 50). The two populations of cells subsequently migrate anteriorly and ventrally until they fuse at the ventral midline to form a single heart field known as the cardiac crescent (18, 48). It is at this point the precursors begin to express the first markers of cardiac development and differentiation, such as the T-box transcription factor *Tbx5*, the homolog of *Drosophila tinman*, *Nkx2.5*, and the transcription factor *mef2* (10, 17, 63). With the fusion of the two precursor populations into the cardiac crescent, the cardiac cells are arranged into the primitive linear heart tube which then undergoes a process of rightward and dorsal looping to bring the future chambers into the correct alignment (22, 36, 69). This initial cardiac crescent, termed the primary heart field, however, has been shown to account for the majority of the total endocardial and myocardial progenitor cells of the adult heart, but not all. Additional populations of cardiac precursors derived from the splanchnic mesoderm and cranial mesoderm, termed the secondary and anterior heartfields, respectively, contribute to the conotruncal myocardium, right ventricle and outflow tract of the developing heart during the process of cardiac looping (27, 35, 66). Following the integration of each of these cardiac progenitor populations, and as embryogenesis proceeds, the cardiac chambers begin to mature and are finally septated to produce the fully functioning heart (6).

Signaling

The complex patterns of cell migration and morphogenesis, which result in the formation of a multi-chambered, beating heart, have been shown to be regulated by a number of secreted factors and signaling pathways originating from the endoderm and the organizer, or node (26, 39, 49). While the exact nature of these secreted signals has yet to be elucidated, a variety of studies have demonstrated roles for molecules such as BMP/TGF-βs, the Wnt antagonists Dkk-1 and Crescent, Notch, and fibroblast growth factors (FGFs) (69). The strongest evidence for a role of BMP molecules in vertebrate heart development comes from studies in chick, frog and mouse. For example, ectopic application of BMP-2 or 4 in chick leads to ectopic expression of cardiac-specific transcription factors such as *Nkx2.5* (52), while activin, a member of the BMP family, has the ability to induce heart tissue in ectodermal explants in *Xenopus laevis* (30). While BMP signals from the endoderm are necessary for proper heart formation, they are not sufficient. Work in chick and frog demonstrates a requirement for the inhibition of canonical Wnt3A and Wnt8 signals in the mesoderm through the diffusion of Dkk and Crescent (33, 51, 65). Additionally, cardiogenesis in *Xenopus* can be stimulated through non-canonical Wnt11 signaling, suggesting the necessity of an intricate network of Wnt signals (42). Notch signaling has also been shown to play an inhibitory role in heart development; however, its actions appear to be downstream of heart field specification. *Xenopus* studies reveal that upregulation of Notch signaling results in an inhibition of myocardial differentiation, while Notch inactivation leads to ectopic expression of myocardial markers (46). A variety of studies have also implied a role for members of the FGF family at numerous stages of cardiac development. The strongest evidence in vertebrates comes from studies of Fgf8 in zebrafish

and mouse. Zebrafish homozygous for mutations in Fgf8 (*acerebellar*, or *ace*) fail to turn on markers of early heart development (45) and disruption of Fgf8 in mouse results in defects in cardiac looping and abnormal cardiovascular patterning (1, 64). Unlike BMP signaling, however, it appears that FGFs are only transiently required.

Transcription factor response to inductive signals

In combination with these inductive signals from the endoderm and organizer, a number of regulatory genes encoding a series of transcription factors are induced and act in the processes of specification, patterning and differentiation of the heart. One of the most well-defined examples, the homeobox gene *Nkx2.5*, a vertebrate homolog of the *Drosophila* gene *tinman*, is expressed in the lateral plate mesoderm coinciding with the time and place of cardiac specification. Genetic studies in a variety of organisms have demonstrated that *Nkx2.5* functions at many stages of cardiac development [reviewed in (21, 23)]. Complete loss of Nkx2.5 in mouse results in early embryonic lethality with severe cardiac defects, while mutant alleles of *Nkx2.5* correlate to inherited atrial-septal defects in humans (31, 32). Recently, Nkx2.5 has been shown to orchestrate a Smad1-dependent negative feedback loop that acts to control heart progenitor specification and proliferation (44). Generally, Nkx2.5 acts to repress BMP2 signaling via Smad1 in the secondary heart field (SHF), while itself is upregulated by active Smad1 signaling in the SHF and foregut endoderm. Active Nkx2.5 in the SHF leads to continued proliferation of cardiac precursors, whereas active BMP signaling through phosphorylated Smad1 leads to the induction of cardiac progenitor genes and a block of proliferation (44). Through this autoregulatory loop, Nkx2.5 is key in the transition

between periods of cardiac induction and progenitor proliferation (44). In addition, it emphasizes the importance of BMP signaling in the proper development of the heart.

Another factor, *Gata4*, encodes a zinc-finger transcription factor essential for cardiogenesis and directly interacts with Nkx2.5 and TBX5 to synergistically activate cardiac gene expression (9, 20). Haploinsufficiency of GATA4 has been linked to congenital septal defects in humans likely due to reduced DNA binding and transactivation of target genes, as well as loss of TBX5 interactions (19, 32, 41, 43).

T-box genes in early cardiogenesis

A large body of evidence supports a role for a number of members of the T-box family of transcription factors in early heart development. Members of the T-box family are characterized by a 180 amino acid, evolutionarily conserved T-box DNA binding domain and have been shown to play important and diverse roles in embryonic development, including germ layer differentiation and organogenesis (47, 57, 68). In support of these family members playing key roles in cardiogenesis, studies have shown these genes to have the ability to respond to mesodermal growth factors such as FGF and TGF-β family members (53, 57). Additionally, mutations in three T-box genes, *Tbx1* (4), *Tbx5* (38), and *Tbx20* (28) have been connected to the clinical manifestation of congenital heart defects in humans (32). Patients presenting with the DiGeorge syndrome phenotype carry either deletions or mutations in the *Tbx1* gene (4, 14), while over 30 different mutations in *Tbx5* have been described in Holt-Oram syndrome patients (24, 38). It was not until recently, however, that one missense and one nonsense mutation were found to be associated with a family history of CHD and a wide spectrum of developmental defects (28). For these reasons, it will be

important to further characterize the roles of these, and other, T-box genes in the complex processes of early cardiogenesis.

Tbx20 and heart development

With *Tbx5* and *Nkx2.5*, *Tbx20* has been shown to be one of the earliest genes expressed in the vertebrate cardiac lineage. *Tbx20* homologs have been described in a variety of organisms including mouse (*Tbx12/20*) (12, 29), zebrafish (*Tbx20/HrT*) (2), chick (26), *Xenopus* (7) and human (34). In *Xenopus*, *Tbx20* is first expressed in gastrula stage embryos in the ventral half of the cement gland. As development proceeds, expression is seen in the migrating heart precursors in concert with the earliest markers of cardiac development, *Nkx2.5* (40) and *Tbx5* (25, 54). Eventually, *Tbx20* can be visualized in all regions of the heart including the atria, ventricles, inflow and outflow tracts and the septum transversum (7, 54). In addition to the heart, *Tbx20* expression is also present in the external jugular vein, lung bud, cloacal aperture, rhombomeres, and a subset of motor neurons (7, 54).

Tbx20 has been shown to be required for proper cardiogenesis in zebrafish, *Xenopus laevis*, mouse and human (8, 11, 28, 56, 59-61). Using morpholino oligonucleotides, Szeto and colleagues demonstrated that zebrafish embryos lacking functional *hrT*, the zebrafish homolog of *Tbx20*, have abnormal hearts and a lack of blood circulation. It appears that early events of heart formation are unaffected, while hearts fail to undergo looping at later stages (60). Studies in *Xenopus*, also using morpholinos, reveal that embryos lacking TBX20 display severe cardiac defects such as an unlooped heart, reduced cardiac mass, and pericardial edema (8). To further the understanding of TBX20 function, multiple groups have gone on to characterize the effects of TBX20 knock-down in the mammalian system.

Embryos with reduced levels of TBX20 by means of transgenic RNA interference or gene targeting are characterized by severe cardiac abnormalities such as outflow tract and right ventricular hypoplasia, disruptions in the early cardiac transcriptional program, and early embryonic lethality (11, 56, 59, 61). In addition, a number of groups have shown that *Tbx20* null embryos exhibit ectopic expression of the T-box gene *Tbx2*, which plays a key role in the repression of chamber differentiation, thus suggesting that TBX20 acts as a transcriptional repressor during early heart development (56, 59). TBX20 has also been shown to interact with a number of other transcription factors in regulating cardiac gene expression. Stennard and colleagues have demonstrated in *Xenopus* that TBX20 interacts with Nkx2.5, GATA4, and GATA5 to synergistically act on cardiac promoters such as *Nppa/ANF* (atrial natriuretic factor) (58). Additionally, Brown *et al*. have shown that TBX20 physically and functionally interacts with TBX5 to control *Xenopus* heart morphogenesis, thus providing the first evidence for direct interactions between T-box family members (8).

DISSERTATION GOALS

Even with these recent studies, the precise regulation, molecular functions or cellular requirements for *Tbx20*, or any one T-box family member, remain elusive. In fact, it is one of the necessities in the field of cardiovascular development to fully dissect the main transcriptional pathways that govern development and to understand how disruptions of the genes that encode components of these pathways lead to aberrations in proper development. To this end, my work focuses on exploring the precise expression and transcriptional regulation of *Tbx20* in the heart and to place it into the larger transcriptional regulatory network that governs early cardiac development. Initially, I focus on a detailed

characterization and comparison of *Tbx20* expression, in addition to the T-box transcription factors *Tbx1*, *Tbx2*, and *Tbx5*, in *X. laevis* and *X. tropicalis*. I go on to describe the identification of a cardiac-specific regulatory element that is necessary for proper *Tbx20* expression, and finally, place *Tbx20* regulation in the context of two major signaling cascades, that of bone morphogenetic protein (BMP) and that of serum response factor (SRF). This is the first description of the regulatory networks controlling the temporal and spatial regulation of *Tbx20*.

In a similar manner, a further understanding of the molecular mechanisms that control cellular proliferation and differentiation is a central theme in developmental biology, and has direct relevance to our studies of cardiac development. Therefore, in addition to my studies examining *Tbx20* regulation, I also describe the first characterization of the transcriptional regulation and differential roles of two co-transcribed microRNAs, miRNA-1 and miRNA-133, in skeletal muscle proliferation and differentiation. This work is a clear demonstration how additional signaling factors, such as miRNAs, can participate in, and have key roles in the transcriptional networks that control early developmental events.

Generally, it is my hope that by fully characterizing the expression and regulation of the *Tbx20* gene in *Xenopus*, and by beginning to define the molecular mechanisms that control the expression of additional regulatory factors, it may be possible to gain a clearer picture of the molecular mechanisms governing heart development and to in turn, better understand the congenital heart defects resulting from aberrations in this process.

REFERENCES

1. **Abu-Issa, R., G. Smyth, I. Smoak, K. Yamamura, and E. Meyers.** 2002. Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. Development **129:**4613-4625.

2. **Ahn, D. G., I. Ruvinsky, A. C. Oates, L. M. Silver, and R. K. Ho.** 2000. *tbx20*, a new vertebrate T-box gene expressed in the cranial motor neurons and developing cardiovascular structures in zebrafish. Mech Dev **95:**253-258.

3. **Antin, P. B., R. G. Taylor, and T. Yatskievych.** 1994. Precardiac mesoderm is specified during gastrulation in quail. Dev Dyn **200:**144-154.

4. **Baldini, A.** 2004. DiGeorge syndrome: an update. Curr Opin Cardiol **19:**201-204.

5. **Blitz, I. L., G. Andelfinger, and M. E. Horb.** 2006. Germ layers to organs: using *Xenopus* to study "later" development. Semin Cell Dev Biol **17:**133-145.

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

7. **Brown, D. D., O. Binder, M. Pagratis, B. A. Parr, and F. L. Conlon.** 2003. Developmental expression of the *Xenopus laevis Tbx20* orthologue. Dev Genes Evol **212:**604-607.

8. **Brown, D. D., S. N. Martz, O. Binder, S. C. Goetz, B. M. J. Price, J. C. Smith, and F. L. Conlon.** 2005. Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development **132:**553-563.

9. **Bruneau, B. G.** 2002. Transcriptional regulation of vertebrate cardiac morphogenesis. Circ Res **90:**509-519.

10. **Bruneau, B. G., M. Logan, N. Davis, T. Levi, C. J. Tabin, J. G. Seidman, and C. E. Seidman.** 1999. Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. Dev Biol **211:**100-108.

11. **Cai, C. L., W. Zhou, L. Yang, L. Bu, Y. Qyang, X. Zhang, X. Li, M. G. Rosenfeld, J. Chen, and S. Evans.** 2005. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development **132**:2475-2487.

12. **Carson, C. T., E. R. Kinzler, and B. A. Parr.** 2000. *Tbx12*, a novel T-box gene, is expressed during early stages of heart and retinal development. Mech Dev **96:**137-140.

13. **Chen, J. N., and M. C. Fishman.** 2000. Genetics of heart development. Trends Genet. **16:**383-388.

14. **Chieffo, C., N. Garvey, W. Gong, B. Roe, G. Zhang, L. Silver, B. S. Emanuel, and M. L. Budarf.** 1997. Isolation and characterization of a gene from the DiGeorge chromosomal region homologous to the mouse *Tbx1* gene. Genomics **43:**267-277.

15. **Cohen-Gould, L., and T. Mikawa.** 1996. The fate diversity of mesodermal cells within the heart field during chicken early embryogenesis. Dev Biol. **177:**265-273.

16. **Dale, L., and J. M. Slack.** 1987. Fate map for the 32-cell stage of *Xenopus laevis*. Development **99:**527-251.

17. **Edmondson, D. G., G. E. Lyons, J. F. Martin, and E. N. Olson.** 1994. *Mef2* gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. Development **120:**1251-1263.

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

19. **Garg, V., I. S. Kathiriya, R. Barnes, M. K. Schluterman, I. N. King, C. A. Butler, C. R. Rothrock, R. S. Eapen, K. Hirayama-Yamada, K. Joo, R. Matsuoka, J. C. Cohen, and D. Srivastava.** 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature **424:**443-447.

20. **Gelb, B. D.** 2004. Genetic basis of congenital heart disease. Curr Opin Cardiol **19:**110-115.

21. **Gruber, P. J., and J. A. Epstein.** 2004. Development gone awry: congenital heart disease. Circ Res **94:**273-283.

22. **Harvey, R. P.** 2002. Patterning the vertebrate heart. Nat Rev Genet **3:**544-556.

23. **Hatcher, C. J., N. Y. Diman, D. A. McDermott, and C. T. Basson.** 2003. Transcription factor cascades in congenital heart malformation. Trends Mol Med **9:**512-515.

24. **Heinritz, W., A. Moschik, A. Kujat, S. Spranger, H. Heilbronner, S. Demuth, A. Bier, M. Tihanyi, S. Mundlos, C. Gruenauer-Kloevekorn, and U. G. Froster.** 2005. Identification of new mutations in the TBX5 gene in patients with Holt-Oram syndrome. Heart **91:**383-384.

25. **Horb, M. E., and G. H. Thomsen.** 1999. Tbx5 is essential for heart development. Development **126:**1739-1751.

26. **Iio, A., M. Koide, K. Hidaka, and T. Morisaki.** 2001. Expression pattern of novel chick T-box gene, *Tbx20*. Dev Genes Evol **211:**559-562.

27. **Kelly, R. G., and M. E. Buckingham.** 2002. The anterior heart-forming field: voyage to the arterial pole of the heart. Trends Genet **18:**210-216.

28. **Kirk, E. P., M. Sunde, M. W. Costa, S. A. Rankin, O. Wolstein, M. L. Castro, T. L. Butler, C. Hyun, G. Guo, R. Otway, J. P. Mackay, L. B. Waddell, A. D. Cole, C. Hayward, A. Keogh, P. Macdonald, L. Griffiths, D. Fatkin, G. F. Sholler, A. M. Zorn, M. P. Feneley, D. S. Winlaw, and R. P. Harvey.** 2007. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. Am J Hum Genet **81:**280-291.

29. **Kraus, F., B. Haenig, and A. Kispert.** 2001. Cloning and expression analysis of the mouse T-box gene *tbx20*. Mech Dev **100:**87-91.

30. **Logan, M., and T. Mohun.** 1993. Induction of cardiac muscle differentiation in isolated animal pole explants of *Xenopus laevis* embryos. Development **118:**865-875.

31. **Lyons, I., L. M. Parsons, L. Hartley, R. Li, J. E. Andrews, L. Robb, and R. P. Harvey.** 1995. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene *Nkx2-5*. Genes Dev **9:**1654-1666.

32. **Mandel, E. M., T.E. Callis, D.Z. Wang, F.L. Conlon.** 2005. Transcriptional Mechanisms of Congenital Heart Disease. Drug Discovery Today: Disease Mechanisms **2:**33-38.

33. **Marvin, M. J., G. Di Rocco, A. Gardiner, S. M. Bush, and A. B. Lassar.** 2001. Inhibition of Wnt activity induces heart formation from posterior mesoderm. Genes Dev **15:**316-327.

34. **Meins, M., D. J. Henderson, S. S. Bhattacharya, and J. C. Sowden.** 2000. Characterization of the human TBX20 gene, a new member of the T-Box gene family closely related to the Drosophila H15 gene. Genomics **67:**317-332.

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

36. **Mohun, T. J., Leong, L.M., Weninger, W. J., and Sparrow, D. B.** 2000. The morphology of heart development in *Xenopus laevis*. Developmental Biology **218:**74-88.

37. **Moody, S. A.** 1987. Fates of the blastomeres of the 32 cell *Xenopus* embryo. Dev. Biol. **122:**300-319.

38. **Mori, A. D., and B. G. Bruneau.** 2004. TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol **19:**211-215.

39. **Nascone, N., and M. Mercola.** 1995. An Inductive role for endoderm in *Xenopus* cardiogenesis. Development **121:**515-523.

40. **Newman, C. S., Krieg, P.A.** 1998. *tinman*-related genes expressed during heart development in *Xenopus*. Genet. **22:**230-238.

41. **Okubo, A., O. Miyoshi, K. Baba, M. Takagi, K. Tsukamoto, A. Kinoshita, K. Yoshiura, T. Kishino, T. Ohta, N. Niikawa, and N. Matsumoto.** 2004. A novel GATA4 mutation completely segregated with atrial septal defect in a large Japanese family. J Med Genet **41:**e97.

42. **Pandur, P., M. Lasche, L. M. Eisenberg, and M. Kuhl.** 2002. Wnt-11 activation of a non-canonical Wnt signaling pathway is required for cardiogenesis. Nature **418:**636-641.

43. **Pehlivan, T., B. R. Pober, M. Brueckner, S. Garrett, R. Slaugh, R. Van Rheeden, D. B. Wilson, M. S. Watson, and A. V. Hing.** 1999. GATA4 haploinsufficiency in patients with interstitial deletion of chromosome region 8p23.1 and congenital heart disease. Am J Med Genet **83:**201-206.

44. **Prall, O. W., M. K. Menon, M. J. Solloway, Y. Watanabe, S. Zaffran, F. Bajolle, C. Biben, J. J. McBride, B. R. Robertson, H. Chaulet, F. A. Stennard, N. Wise, D. Schaft, O. Wolstein, M. B. Furtado, H. Shiratori, K. R. Chien, H. Hamada, B. L. Black, Y. Saga, E. J. Robertson, M. E. Buckingham, and R. P. Harvey.** 2007. An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation. Cell **128:**947-959.

45. **Reifers, F., E. C. Walsh, S. Leger, D. Y. Stainier, and M. Brand.** 2000. Induction and differentiation of the zebrafish heart requires fibroblast growth factor 8 (fgf8/acerebellar). Development **127:**225-235.

46. **Rones, M. S., K. A. McLaughlin, M. Raffin, and M. Mercola.** 2000. Serrate and Notch specify cell fates in the heart field by suppressing cardiomyogenesis. Development **127:**3865-3876.

47. **Ryan, K., and A. J. Chin.** 2003. T-box genes and cardiac development. Birth Defects Res Part C Embryo Today **69:**25-37.

48. **Sater, A. K., and A. G. Jacobson.** 1990. The restriction of the heart morphogenetic field in *Xenopus laevis*. Dev. Biol. **140:**328-336.

49. **Sater, A. K., and A. G. Jacobson.** 1990. The role of the dorsal lip in the induction of heart mesoderm in *Xenopus laevis*. Mol Cell Endocrinol **70:**109-116.

50. **Sater, A. K., and A. G. Jacobson.** 1989. The specification of heart mesoderm occurs during gastrulation in *Xenopus laevis*. Development **106:**519-529.

51. **Schneider, V. A., and M. Mercola.** 2001. Wnt antagonism initiates cardiogenesis in *Xenopus laevis*. Genes Dev **15:**304-315.

52. **Schultheiss, T. M., J. B. Burch, and A. B. Lassar.** 1997. A role for bone morphogenetic proteins in the induction of cardiac myogenesis. Genes Dev **11:**451-462.

53. **Showell, C., O. Binder, and F. L. Conlon.** 2004. T-box genes in early embryogenesis. Dev Dyn **229:**201-218.

54. **Showell, C., K. S. Christine, E. M. Mandel, and F. L. Conlon.** 2006. Developmental expression patterns of *Tbx1*, *Tbx2*, *Tbx5*, and *Tbx20* in *Xenopus tropicalis*. Dev Dyn **235:**1623-1630.

55. **Showell, C., and F. L. Conlon.** 2007. Decoding development in *Xenopus tropicalis*. Genesis **45:**418-426.

56. **Singh, M. K., V. M. Christoffels, J. M. Dias, M. O. Trowe, M. Petry, K. Schuster-Gossler, A. Burger, J. Ericson, and A. Kispert.** 2005. Tbx20 is essential for cardiac chamber differentiation and repression of *Tbx2*. Development **132:**2697-2707.

57. **Smith, J.** 1999. T-box genes: what they do and how they do it. Trends Genet. **15:**154-158.

58. **Stennard, F. A., M. W. Costa, D. A. Elliott, S. Rankin, S. J. Haast, D. Lai, L. P. McDonald, K. Niederreither, P. Dolle, B. G. Bruneau, A. M. Zorn, and R. P. Harvey.** 2003. Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart. Dev Biol **262:**206-24.

59. **Stennard, F. A., M. W. Costa, D. Lai, C. Biben, M. B. Furtado, M. J. Solloway, D. J. McCulley, C. Leimena, J. I. Preis, S. L. Dunwoodie, D. E. Elliott, O. W. Prall, B. L. Black, D. Fatkin, and R. P. Harvey.** 2005. Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. Development **132**:2451-2462.

60. **Szeto, D. P., K. J. Griffin, and D. Kimelman.** 2002. HrT is required for cardiovascular development in zebrafish. Development **129:**5093-5101.

61. **Takeuchi, J. K., M. Mileikovskaia, K. Koshiba-Takeuchi, A. B. Heidt, A. D. Mori, E. P. Arruda, M. Gertsenstein, R. Georges, L. Davidson, R. Mo, C. C. Hui, R. M. Henkelman, M. Nemer, B. L. Black, A. Nagy, and B. G. Bruneau.** 2005. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. Development **132**:2463-2474.

62. **Tam, P. P., M. Parameswaran, S. J. Kinder, and R. P. Weinberger.** 1997. The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development **124:**1631-1642.

63. **Tonissen, K. F., T. A. Drysdale, T. J. Lints, R. P. Harvey, and P. A. Krieg.** 1994. *XNkx-2.5*, a *Xenopus* gene related to *Nkx-2.5* and *tinman*: evidence for a conserved role in cardiac development. Dev. Biol. **162:**325-328.

64. **Trumpp, A., M. J. Depew, J. L. Rubenstein, J. M. Bishop, and G. R. Martin.** 1999. Cre-mediated gene inactivation demonstrates that FGF8 is required for cell survival and patterning of the first branchial arch. Genes Dev. **13:**3136-3148.

65. **Tzahor, E., and A. B. Lassar.** 2001. Wnt signals from the neural tube block ectopic cardiogenesis. Genes Dev **15:**255-260.

66. **Waldo, K. L., D. H. Kumiski, K. T. Wallis, H. A. Stadt, M. R. Hutson, D. H. Platt, and M. L. Kirby.** 2001. Conotruncal myocardium arises from a secondary heart field. Development **128:**3179-3188.

67. **Warkman, A. S., and P. A. Krieg.** 2007. *Xenopus* as a model system for vertebrate heart development. Semin Cell Dev Biol **18:**46-53.

68. **Wilson, V., and F. L. Conlon.** 2002. The T-box family. Genome Biol **3:**REVIEWS3008.

69. **Zaffran, S., and M. Frasch.** 2002. Early signals in cardiac development. Circ Res **91:**457-469.

CHAPTER 2

TRANSCRIPTIONAL MECHANISMS OF CONGENITAL HEART DISEASE

PREFACE

Chapter 2 focuses on the current understanding of the transcription factors that have been linked to familial congenital heart diseases. This work was published, in part, as a review article, and represents a first-author manuscript and collaboration with Thomas E. Callis and Da-Zhi Wang. At the time of publication, mutations or deletions of the transcription factors *Tbx1*, *Nkx2.5*, *Gata4*, *Tbx5*, *Sall1*, *TFAP2B*, and *Eya4* had been associated with human disease. Since that time, however, missense and nonsense mutations in the cardiac transcription factor *Tbx20* have been linked to a variety of congenital heart defects. This emphasizes the rapid pace at which the fields of cardiac development and disease are advancing, and further supports the need for continuing research in these areas.

Mandel, E.M., T.E. Callis, D.Z. Wang, and F.L. Conlon. 2005. Transcriptional mechanisms of congenital heart disease. Drug Discovery Today: Disease Mechanisms **2(1)**:33-38.

SUMMARY

Over the past decade, clinical studies have identified a number of congenital heart diseases associated with mutations in cardiac transcription factors. Recent reports have shown that several of these transcription factors physically interact with one another, implying they may act in similar molecular pathways. Here, we outline familial heart diseases linked to cardiac transcription factors and describe some of the emerging technologies being developed as potential therapeutics for these diseases.

INTRODUCTION

Congenital heart defects (CHDs) threaten nearly 1% of all newborns and pose a significant threat of infant death; however, the underlying genetic mechanisms of many CHDs remain elusive. Most likely, the majority of these defects have a basis in the complex process of cardiogenesis. Heart development involves a series of highly coordinated events including cell proliferation, differentiation, migration and morphogenesis, and a number of genes involved in these processes have been identified as potential causes of specific cardiac anomalies. Transcription factors are major regulators of developmental processes and play essential roles in cardiogenesis (Fig. 2.1). Here, we review 6 CHDs associated with deletions or mutations of transcription factors and the current understanding of their molecular bases.

DiGeorge Syndrome: *Tbx1*

The presentation of DiGeorge syndrome (OMIM #188400) is one consequence of the most common human genetic deletion, monoallelic microdeletion (see glossary) of chromosome 22q11.2. In most cases, the heterozygous deletion eliminates approximately

3Mbp of the long arm of chromosome 22, resulting in the loss of an estimated 30 genes (3). Recent studies have described highly variable clinical indications of patients with chromosome 22q11.2 deletions, even within the same pedigree. However, CHDs are the most common feature of DiGeorge syndrome, or *del22q11*, and may include tetralogy of Fallot, interruption of the aortic arch type B, ventricular septal defects, pulmonary atresia, or persistent truncus ateriosus (see glossary for terms; Fig. 2.2) (3, 51).

The use of mouse genetics has recently clarified which of the numerous genes deleted in *del22q11* may be responsible for the DiGeorge syndrome phenotypes. Targeted mutations in the mouse genome have allowed the majority of the DiGeorge syndrome clinical manifestations to be attributed to haploinsufficiency of *Tbx1* (**AF012130**), one of the genes deleted in *del22q11* patients (3, 15). *Tbx1* is a member of the T-box family of transcription factors and is involved in the patterning of the pharyngeal endoderm and aortic arches, as well as cardiac outflow tract development in a gene dosage-dependent manner (51). Attempts to further connect *Tbx1* to DiGeorge syndrome have led to searches for mutations in this gene in patients lacking the typical chromosomal deletion. To this end, five patients have been identified as carrying only a *Tbx1* gene mutation (50). Though these individuals do not exhibit all characteristics of DiGeorge syndrome, this demonstrates that mutations in human *Tbx1*, as in mouse, are capable of causing many of the defects associated with *del22q11*. In an effort to understand the molecular mechanisms of TBX1 function, recent observations have resulted in a model in which *Fgf8* in the pharyngeal endoderm is regulated by TBX1 to control the proper patterning of the aortic arch through epithelial-mesenchymal interactions (Table 2.1) (51). Additionally, *Tbx1* transcription has been shown to be regulated by the sonic hedgehog (shh) signaling pathway via the *Foxc1* and *Foxc2*

transcription factors which are expressed in the head mesenchyme and the mesenchyme surrounding the aortic arch arteries (Table 2.1) (51). Together, studies such as these demonstrate signaling cascades by which *Tbx1* is transcribed and may initiate proper patterning events; however, the complete mechanism of TBX1 action remains unknown.

Familial Cardiac Septal Defects: *Nkx2.5, Gata4* and *Tbx20*

Cardiac septal defects (CSDs) are a common form of CHD and are defined by a hole in the septal wall allowing blood transfer between the atria or ventricles. Atrial septal defects (ASDs) affect over one in 1000 live births, while ventricular septal defects (VSDs) are the most prevalent CHD, occurring in approximately one in 300 live births (Fig. 2.2). Over time, persistent left-to-right shunting of blood between the atria or ventricles leads to pulmonary hypertension, arrhythmias, and atrial and ventricular dysfunction. Fortunately, severe ASDs and VSDs can be treated by surgical- or catheter-based procedures that employ a prosthetic patch to close the defect. Despite the high incidence of CSDs, the precise molecular mechanisms directing septal morphogenesis remain unclear. However, genetic studies have implicated mutations in the *Nkx2.5* and *Gata4* loci as genetic causes of familial CSDs.

Mutant alleles of the *Nkx2.5* locus (**BC025711**) correlate with ASDs in rare families in which the defect is inherited (Fig. 2.2) (44). Genetic studies in a wide variety of organisms demonstrate that NKX2.5 functions at many stages of cardiac development and in a variety of cardiac tissues (18, 19). Complete loss of NKX2.5 in mice results in early embryonic lethality with severe cardiac defects (30), while mice heterozygous for the *Nkx2.5* allele only occasionally suffer ASDs (5). This suggests that genetic modifiers are important for ASD penetrance. NKX2.5 interacts with other transcription factors associated with

CHDs such as GATA4 and TBX5, and many cardiac-specific genes contain NKX2.5 binding sites in their promoters, highlighting the importance of *Nkx2.5* in the cardiac transcriptional program (Table 2.1) (12). Chien *et al.* (2004) reported that mice harboring a ventricular muscle-cell restricted knockout of *Nkx2.5* mimic CHD and implicated persistent BMP-10 (Bone Morphogenetic Protein-10) expression as playing an important role in the onset and progression of observed cardiac defects (39). This study suggests that antagonizing BMP-10 signals could represent a new therapeutic approach to prevent progression of *Nkx2.5*-associated CHDs.

ASDs, as well as VSDs and atrioventricular septal defects, are also associated with *Gata4* (**AY740706**) haploinsufficiency (Fig. 2.2) (16, 36, 40). A study of a large pedigree revealed a missense mutation in *Gata4* linked to an autosomal dominant disorder where ASD was fully penetrant. *Gata4* encodes a zinc-finger transcription factor essential for cardiogenesis, and directly interacts with the cardiac transcription factors NKX2.5 and TBX5 to synergistically activate cardiac gene expression (12, 17). Inherited mutations in *Gata4* result in reduced DNA binding and transactivation of target genes, as well as loss of TBX5 interaction (16). In addition to NKX2.5 and TBX5, GATA4 associates with a variety of binding partners thought to create specific transcriptional complexes that confer tissue-specific gene expression during heart development (Table 2.1) (12, 19).

Familial defects in cardiac septation have recently been associated with a third cardiac transcription factor, *Tbx20* (**NM_020496**). As previously described, *Tbx20* is a member of the T-box family of transcription factors and is one of the earliest markers of vertebrate cardiogenesis (1, 10, 14, 22, 27, 33). TBX20 has been shown to be required for proper vertebrate heart development, as reduced levels of TBX20 have been shown to disrupt

cell specification, proliferation and chamber differentiation (13, 45, 47, 48). Recently, a novel screen of 352 CHD-affected propositi for mutations in the *Tbx20* locus revealed one missense and one nonsense mutation to be associated with a family history of CHD (24). The missense mutation, I152M, which disrupts a highly conserved amino acid of the T-box DNA-binding domain, results in decreased *Tbx20* function and is clinically manifest by both ASDs and VSDs (24). The nonsense mutation, Q195X, on the other hand, results in a truncated TBX20 protein product through the introduction of a stop codon into one exon of the region encoding the T-box domain. Probands identified to carry the Q195X nonsense mutation presented with a wider range of CHDs which includes not only ASDs and VSDs, but also mitral valve stenosis, DCM, apicolateral hypertrophy, dilated LV and pulmonary hypertension (24). This broad range of defects, taken in the context of the nature of the specific mutation is in part representative of TBX20 haploinsufficiency and is consistent with the wide expression profile of *Tbx20* and the severe phenotypes observed in TBX20 null mice (13, 45, 47, 48). It is interesting, however, that the human clinical presentation resulting from the Q195X mutation is much more severe than the phenotype of heterozygous mice. This is similar to mouse models of *Nkx2.5* and may suggest an importance of modifier genes in TBX20 expression (12).

Holt-Oram Syndrome: *Tbx5*

Holt-Oram syndrome (HOS, OMIM #142900) is an autosomal dominant condition that occurs in approximately one of every 100,000 live births. HOS generally presents highly variable phenotypes including both upper limb and cardiac defects. Though rare, there is much to learn from its presentation of CHDs, which range from single or multiple ASDs and

VSDs, to more complex malformations such as tetralogy of Fallot and hypoplastic left heart syndrome (see glossary; Fig. 2.2) (34). Mild to severe cardiac arrhythmias are also common (4).

The genomic locus responsible for HOS phenotypes was previously mapped to chromosome 12q24.1. Since then, HOS has been linked to more than 30 mutations distributed throughout *Tbx5* (**U80987**), generally thought to result in *Tbx5* haploinsufficiency (21, 34). *Tbx5*, like *Tbx1*, is a T-box containing transcription factor that is essential for proper vertebrate tissue patterning and differentiation (37). Though familial studies and studies in mouse have attempted to correlate the location of the many *Tbx5* mutations along the locus with the wide variation in severity of limb and cardiac defects, there is currently insufficient evidence to support such a hypothesis (34). Currently, it is thought that loss of transactivation, reduced interaction with other cardiac transcription factors such as NKX2.5, GATA4, and TBX20, or mis-sorting of mutant forms of TBX5 are main causes for HOS pathogenesis (Table 2.1) (11, 17, 19).

Okihiro Syndrome: *Sall4*

Okihiro Syndrome (OMIM #126800) is an autosomal dominant condition consisting of Duane anomaly, radial ray defects and deafness (see glossary). The phenotype may include cardiac defects, anal stenosis, pigmentary disturbance, renal abnormalities, or facial asymmetries. The specific cardiac defects are most often ASDs, VSDs, or tetralogy of Fallot (Fig. 2.2) (8, 26).

Familial studies of individuals affected by Okihiro syndrome have identified mutations in the *Sall4* gene (**NM_020436**) and suggest that haploinsufficiency of this gene is

responsible for the clinical phenotype (2, 7, 25). *Sall4* (*spalt-like 4*) is a member of the *Sal* gene family, which encodes a group of four probable zinc-finger transcription factors (28). Thus far, a total of 11 different mutations over the entire *Sall4* gene have been described in relation to Okihiro syndrome (2, 25, 26). In addition, Borozdin (2004) and colleagues demonstrated that Okihiro syndrome can also be caused by deletions of either the entire *Sall4* gene or of single coding exons (7). Based on work with the closely related *Sall1*, it is likely that these mutations result in truncated proteins, possibly having the dominant effect of an upregulated repressor (23). However, at this point there are no known upstream effectors or downstream targets of SALL4.

Char Syndrome: *TFAP2B*

Char syndrome (OMIM #169100) is an autosomal dominant trait characterized by facial dysmorphism, hand anomalies, and patent ductus arteriosus (see glossary; Fig. 2.2). Char syndrome has been mapped to chromosome 6p12-p21 and further analyses point to inherited mutations within the *TFAP2B* (transcription factor AP-2 beta) (**NM_003221**) locus as a genetic cause of Char syndrome (31, 42, 52). TFAP2B encodes a neural crest-related transcription factor belonging to the TFAP family, whose members play an important role in retinoic acid-induced differentiation (29). Char syndrome likely results from abnormal neural crest development, as neural crest cells are important for the development of several affected tissues (35). TFAP2B mutations associated with Char syndrome inhibit target gene activation through a dominant-negative mechanism or cause abnormal mRNA splicing resulting in TFAP2B haploinsufficiency (31, 42). However, the precise molecular

mechanisms underlying the effects of aberrant TFAP2B activity resulting in Char syndrome remain to be elucidated.

Dilated Cardiomyopathy with Sensorineural Hearing Loss: *Eya4*

Cardiomyopathy is the leading cause of heart failure and is most commonly associated with a dilated cardiomyopathy (DCM) phenotype, defined by increased diastolic and systolic ventricular volumes and contractile dysfunction (41). Often, DCM is presented in conjunction with defects of the inner ear resulting in sensorineural hearing loss (SNHL). While not typically characterized as a classical CHD, cardiomyopathy is discussed here as the origins of its pathophysiology may also result from mutations in cardiac transcription factors.

Several studies have demonstrated that approximately 25% to 30% of DCM cases may be familial (41). Until recently, the significant mortality and late onset of this disease hindered work to identify the genomic location of the responsible disease loci. Schönberger and colleagues have identified a human mutation that causes dilated cardiomyopathy and associated heart failure in addition to previously described sensorineural hearing loss (43, 49). The identified mutation is a 4846 bp deletion of the human gene *Eya4* (**Y17114**), one of four vertebrate orthologs of the *Drosophila melanogaster* gene *eyes absent* (*eya*) (9). EYA4 is a transcriptional coactivator that interacts with members of the *sine oculis* family (*Six1-Six6*) and *Dach* transcription factors to lead to gene activation (Table 2.1) (6, 9, 20). The characterization of the human mutation is supported by work in zebrafish, as attenuated *eya4* levels produce the morphological and hemodynamic features of heart failure. In addition,

Schönberger *et al.* (2002) demonstrate critical roles for EYA4-SIX regulation of transcription in normal heart function (43).

Potential Therapies

In general, transcription factors have historically been poor targets of drug therapy due to their nuclear localization, lack of enzymatic activity, and the difficulty associated with reprogramming transcriptional networks (Fig. 2.1). Presently, the most effective therapy for cardiac diseases is heart transplantation. However, due to the shortage of organs, cost, and inaccessibility of treatment for most affected individuals, this remains a limited therapeutic option. Alternative treatment is the administration of drugs that improve myocardial contractility, though this treatment is only effective as a short-term therapy, with the 5-year survival rate using current agents being less then 60% (32). More recently, new strategies have focused on two main approaches for treatment of transcription factor-associated heart disease, cardiac stem cell transplants and chemical modulators of transcriptional activity.

The ability to isolate and propagate cell populations that can differentiate into cardiomyocytes *in vivo* offers the opportunity to treat a wide range of cardiac diseases. The existence of cardiac precursor or stem cells in adults remains a contentious issue. However, recent reports suggest that cardiac precursor or stem cells are present, albeit, in a very low number. In addition to endogenous cardiac stem cells, other studies have shown that multipotential cells, most notably embryonic stem (ES) cells and bone marrow-derived stem cells have, under defined conditions, differentiated into cardiomyocytes. Although these studies offer the promise of growing cells for use in repairing damaged cardiac tissue, three major hurdles must be overcome before stem cells can be considered as a therapy for cardiac

disease. First, the molecular, biochemical, and cellular properties of these different cell populations must be established. Second, studies must demonstrate that precursor cell populations can be maintained and expanded to suitable numbers to be used as a cardiac therapy while maintaining their multipotentiality. Finally, results must show multipotential cells, once transplanted to the heart, can give rise to functioning cardiomyocytes while not undergoing uncontrolled differentiation leading to cardiac teratomas or fibrosis (38).

An alternative therapeutic strategy for transcription factor-associated cardiac disease is to screen small chemical libraries to identify agents that either exacerbate or ameliorate transcription factor activities (Fig. 2.1). These agents could either act in an intercellular signaling cascade that turns on, off, or modifies transcription factor activities, most notably agents that act in the calcium or phosphate signaling pathways, or act directly on transcription co-factors such as histone acetyltransferases (HATs) or histone deactylases (HDACs) (32). The major obstacle is the availability of an inexpensive, quick screen for these molecules. However, recent observations showing the sequence, expression, and function of many cardiac disease-associated transcription factors are evolutionarily conserved open the possibility of using fish or frog model systems as bioassays to test for agents that modulate these pathways.

SUMMARY AND CONCLUSIONS

Given the number of transcription factors demonstrated to play essential roles in vertebrate cardiogenesis, the current pool of CHD-associated transcription factors is likely underrepresented. It is our speculation that more correlations between mutations in cardiac transcription factors and CHD will be made. To this end, disruptions in the function of at

least six cardiac transcription factors have been associated with human CHD. Haploinsufficiency of the genes *Tbx1*, *Tbx5*, *GATA4*, and *Sall4* have been correlated with CHDs such as DiGeorge syndrome, Holt-Oram syndrome, familial ASDs and VSDs and Okihiro syndrome while similar disruptions in *Nkx2.5*, *TFAP2B*, and *Eya4* are associated with familial ASDs, Char syndrome and cardiomyopathies. While the identification of genes associated with CHD is an important first step towards the goal of curing cardiovascular disease, it has become clear that understanding the genetic pathways and the molecular mechanisms of transcription factors will be key to our ability to identify therapeutic agents for CHD. Based on our current understanding of these mechanisms and of heart development in general, possible treatment options may eventually grow to include cardiac stem cell transplants and chemical agents. However, these possibilities lie in the future, and their development will rely upon studies using a variety of animal models and our growing knowledge of CHD.

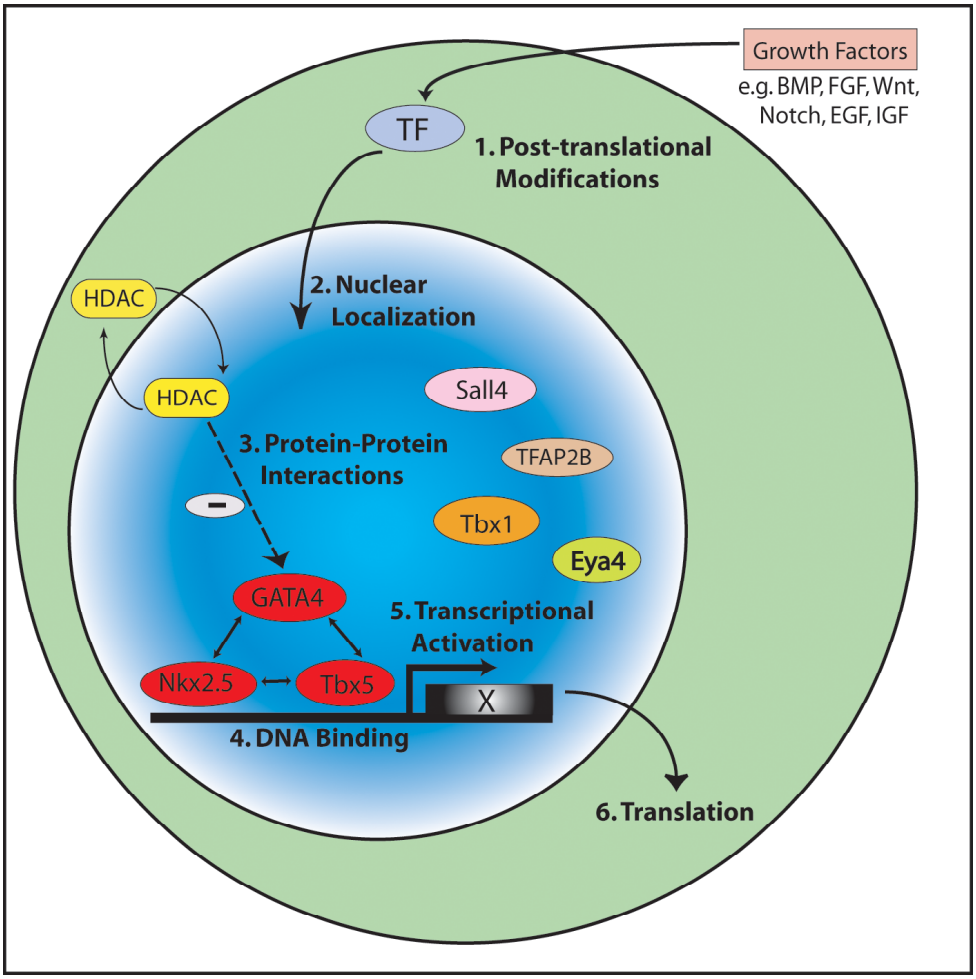


Figure 2.1 General pathways of cardiac transcription factor regulation. 1. Post-translational modification of cardiac transcription factors (TF) in response to growth factor signaling pathways. 2. Nuclear localization of the cardiac TFs. 3. Protein-protein interactions of cardiac TF including potential association with the transcriptional repressors, histone deacetylases (HDAC). 4. Sequence specificity or affinity of DNA binding. 5. Regulation of transcriptional activity including cardiac TF-associated activation or repression. 6. Nuclear export and translation of downstream target genes. Each step represents a potential target for drug therapies. *BMP* – bone morphogenetic protein, *FGF* – fibroblast growth factor, *EGF* – epidermal growth factor, *IGF* – insulin-like growth factor, *TF* – transcription factor, *HDAC* – histone deacetylase

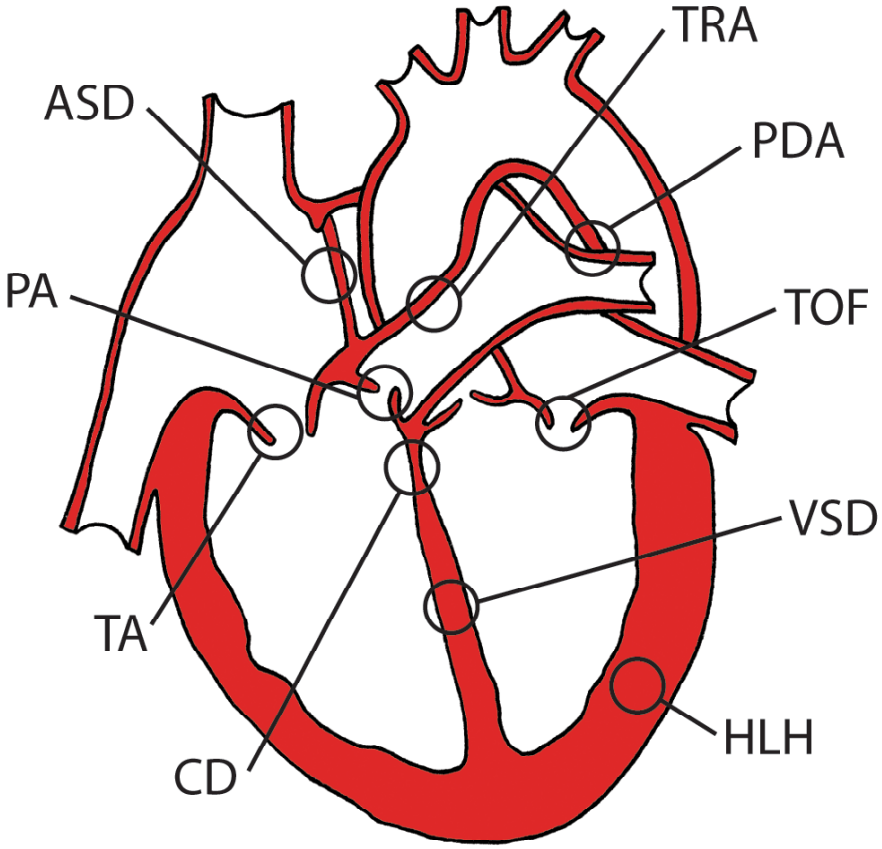


Figure 2.2. Congenital heart defects associated with mutations in cardiac transcription factors. Mutations in cardiac transcription factors such as *Tbx1*, *Nkx2.5*, *Gata4*, *Tbx5*, *Sall4*, and *TFAP2B* have been associated with multiple human congenital heart defects including atrial septal defects (ASD), ventricular septal defects (VSD), tetralogy of Fallot (TOF), conduction system defects (CD), hypoplastic left heart (HLH), pulmonary atresia (PA), patent ductus arteriosus (PDA), tricuspid atresia (TA), and truncus arteriosus (TRA). Schematic shows the relative position of each defect.

Human Clinical Manifestation	**Associated Transcription Factor**	**Co-factors/ Upstream Molecules**	**Potential Downstream Cardiac Genes**	**References**
DiGeorge Syndrome	*Tbx1*	VEGF, FOXc1, FOXc2	*Fgf8*	(3, 12, 15, 18, 19, 50, 51)
Familial ASD	*Nkx2.5*	GATA4, TBX5	*Nppa/ANF, Bnp, eHand, Mef2C Mlc2V, Msx2, N-Myc*	(5, 12, 17-19, 30, 44)
Familial ASD/VSD	*Gata4*	FOG2, GATA6, MEF2C, NFATc4, NKX2.5, SRF, TBX5	*Nppa/ANF,* a/b-*MHC, Cardiac* a-*actin, Cardiac TnC,Cardiac TnI, Gata6, Nkx2.5*	(12, 16-19, 36, 40)
Septal and Valvulogenesis Defects and Cardiomyopathy	*Tbx20*	GATA4, GATA5, NKX2.5, TBX5	*Nppa/ANF, Tbx2*	(11, 45-47)
Holt-Oram Syndrome	*Tbx5*	GATA4, NKX2.5, TBX20	*Nppa/ANF, Cx40, Gata4, Hey2, Irx4, Mlc2v, Nkx2.5*	(4, 11, 12, 17-19, 21, 34, 37)
Okihiro Syndrome	*Sall4*	Unknown	Unknown	(2, 7, 8, 19, 23, 25, 26, 28)
Char Syndrome	*TFAP2B*	Unknown	Unknown	(17, 31, 42, 52)
Dilated Cardiomyopathy with Sensorineural Hearing Loss	*Eya4*	SIX, DACH	Unknown	(6, 9, 19, 20, 43, 49)

Table 2.1. Human CHDs, associated transcription factors and molecular interactions

Abbreviations: VEGF-vascular endothelial growth factor; FOXc1/2-human forkhead-box subfamily c1/2; GATA-GATA binding protein; TBX-T-box; FOG2-friend of GATA; MEF2C-myocyte enhancer factor 2C isoform; NFATc4-nuclear factor of activated T cells, cytoplasmic, calcineurin-dependent-4; NKX2.5-NK2-related homeobox; SRF-serum response factor; *Fgf8*-fibroblast growth factor 8; *Nppa/ANF*-natriuretic peptide precursor a/atrial natriuretic factor; *Bnp*-brain natriuretic peptide; *eHAND*-heart- and neural crest derivatives-expressed 1; *mlc2v*-myosin light chain-2 ventricular isoform; *Msx2*-muscle segment homeobox 2; *N-Myc*-neuroblastoma-myelocytomatosis viral-related oncogene; α/β-*MHC*-alpha-, beta- myosin heavy chain; *TnC*-troponin C; *TnI*-troponin I; *Cx40*-connexin 40; *Hey2*-hairy/enhancer of split-related with YRPW motif-2; *Irx4*-iroquois 4

GLOSSARY

Monoallelic microdeletion – Loss of a small fragment of one allele of a gene

Tetralogy of Fallot – A congenital heart defect characterized by a malformation of the ventricular septum, enlargement of the right ventricle, misplacement of the aortic origin, and a narrowing of the pulmonary artery

Interruption of the aortic arch type B – A congenital heart defect characterized by an interruption between the left carotid artery and the left subclavian artery

Ventricular Septal Defect – A congenital heart defect characterized by an abnormal opening between the cardiac ventricles, usually resulting from failure of the spiral septum to close the interventricular foramen and allowing blood to pass directly from the left to the right ventricle

Pulmonary Atresia – A congenital heart defect characterized by the absence of the normal valve into the pulmonary artery

Persistent Truncus Arteriosus – A congenital cardiovascular defect resulting from the failure of the septum between the aorta and pulmonary trunk to develop and characterized by a common arterial trunk opening out of both ventricles with the pulmonary arteries branching from the ascending common trunk

Hypoplastic Left Heart Syndrome – A congenital heart defect in which the left side of the heart is underdeveloped resulting in insufficient blood flow

Duane Anomaly – A congenital eye movement disorder characterized by an absence of abduction, restricted adduction and retraction of the globe on attempted adduction

Radial Ray Defects – Congenital malformations of the forelimbs

Patent Ductus Arteriosus – A cardiovascular defect in which the ductus arteriosus fails to close after birth

REFERENCES

1. **Ahn, D. G., I. Ruvinsky, A. C. Oates, L. M. Silver, and R. K. Ho.** 2000. *tbx20*, a new vertebrate T-box gene expressed in the cranial motor neurons and developing cardiovascular structures in zebrafish. Mech Dev **95:**253-258.

2. **Al-Baradie, R., K. Yamada, C. St Hilaire, W. M. Chan, C. Andrews, N. McIntosh, M. Nakano, E. J. Martonyi, W. R. Raymond, S. Okumura, M. M. Okihiro, and E. C. Engle.** 2002. Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. Am J Hum Genet **71:**1195-1199.

3. **Baldini, A.** 2004. DiGeorge syndrome: an update. Curr Opin Cardiol **19:**201-204.

4. **Basson, C. T., G. S. Cowley, S. D. Solomon, B. Weissman, A. K. Poznanski, T. A. Traill, J. G. Seidman, and C. E. Seidman.** 1994. The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). N Engl J Med **330:**885-891.

5. **Biben, C., R. Weber, S. Kesteven, E. Stanley, L. McDonald, D. A. Elliott, L. Barnett, F. Koentgen, L. Robb, M. Feneley, and R. P. Harvey.** 2000. Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene *Nkx2-5*. Circ Res **87:**888-895.

6. **Bonini, N. M., W. M. Leiserson, and S. Benzer.** 1993. The eyes absent gene: genetic control of cell survival and differentiation in the developing Drosophila eye. Cell **72:**379-395.

7. **Borozdin, W., D. Boehm, M. Leipoldt, C. Wilhelm, W. Reardon, J. Clayton-Smith, K. Becker, H. Muhlendyck, R. Winter, O. Giray, F. Silan, and J. Kohlhase.** 2004. SALL4 deletions are a common cause of Okihiro and acro-renal-ocular syndromes and confirm haploinsufficiency as the pathogenic mechanism. J Med Genet **41:**e113.

8. **Borozdin, W., M. J. Wright, R. C. Hennekam, M. C. Hannibal, Y. J. Crow, T. E. Neumann, and J. Kohlhase.** 2004. Novel mutations in the gene SALL4 provide further evidence for acro-renal-ocular and Okihiro syndromes being allelic entities, and extend the phenotypic spectrum. J Med Genet **41:**e102.

9. **Borsani, G., A. DeGrandi, A. Ballabio, A. Bulfone, L. Bernard, S. Banfi, C. Gattuso, M. Mariani, M. Dixon, D. Donnai, K. Metcalfe, R. Winter, M. Robertson, R. Axton, A. Brown, V. van Heyningen, and I. Hanson.** 1999. EYA4, a novel vertebrate gene related to Drosophila eyes absent. Hum Mol Genet **8:**11-23.

10. **Brown, D. D., O. Binder, M. Pagratis, B. A. Parr, and F. L. Conlon.** 2003. Developmental expression of the *Xenopus laevis Tbx20* orthologue. Dev Genes Evol **212:**604-607.

11. **Brown, D. D., S. N. Martz, O. Binder, S. C. Goetz, B. M. J. Price, J. C. Smith, and F. L. Conlon.** 2005. Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development **132:**553-563.

12. **Bruneau, B. G.** 2002. Transcriptional regulation of vertebrate cardiac morphogenesis. Circ Res **90:**509-519.

13. **Cai, C. L., W. Zhou, L. Yang, L. Bu, Y. Qyang, X. Zhang, X. Li, M. G. Rosenfeld, J. Chen, and S. Evans.** 2005. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development **132**:2475-2487

14. **Carson, C. T., E. R. Kinzler, and B. A. Parr.** 2000. *Tbx12*, a novel T-box gene, is expressed during early stages of heart and retinal development. Mech Dev **96:**137-140.

15. **Chieffo, C., N. Garvey, W. Gong, B. Roe, G. Zhang, L. Silver, B. S. Emanuel, and M. L. Budarf.** 1997. Isolation and characterization of a gene from the DiGeorge chromosomal region homologous to the mouse *Tbx1* gene. Genomics **43:**267-277.

16. **Garg, V., I. S. Kathiriya, R. Barnes, M. K. Schluterman, I. N. King, C. A. Butler, C. R. Rothrock, R. S. Eapen, K. Hirayama-Yamada, K. Joo, R. Matsuoka, J. C. Cohen, and D. Srivastava.** 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature **424:**443-447.

17. **Gelb, B. D.** 2004. Genetic basis of congenital heart disease. Curr Opin Cardiol **19:**110-5.

18. **Gruber, P. J., and J. A. Epstein.** 2004. Development gone awry: congenital heart disease. Circ Res **94:**273-283.

19. **Hatcher, C. J., N. Y. Diman, D. A. McDermott, and C. T. Basson.** 2003. Transcription factor cascades in congenital heart malformation. Trends Mol Med **9:**512-515.

20. **Heanue, T. A., R. Reshef, R. J. Davis, G. Mardon, G. Oliver, S. Tomarev, A. B. Lassar, and C. J. Tabin.** 1999. Synergistic regulation of vertebrate muscle development by *Dach2*, *Eya2*, and *Six1*, homologs of genes required for Drosophila eye formation. Genes Dev **13:**3231-3243.

21. **Heinritz, W., A. Moschik, A. Kujat, S. Spranger, H. Heilbronner, S. Demuth, A. Bier, M. Tihanyi, S. Mundlos, C. Gruenauer-Kloevekorn, and U. G. Froster.** 2005. Identification of new mutations in the TBX5 gene in patients with Holt-Oram syndrome. Heart **91:**383-384.

22. **Iio, A., M. Koide, K. Hidaka, and T. Morisaki.** 2001. Expression pattern of novel chick T-box gene, *Tbx20*. Dev Genes Evol **211:**559-562.

23. **Kiefer, S. M., B. W. McDill, J. Yang, and M. Rauchman.** 2002. Murine *Sall1* represses transcription by recruiting a histone deacetylase complex. J Biol Chem **277:**14869-14876.

24. **Kirk, E. P., M. Sunde, M. W. Costa, S. A. Rankin, O. Wolstein, M. L. Castro, T. L. Butler, C. Hyun, G. Guo, R. Otway, J. P. Mackay, L. B. Waddell, A. D. Cole, C. Hayward, A. Keogh, P. Macdonald, L. Griffiths, D. Fatkin, G. F. Sholler, A. M. Zorn, M. P. Feneley, D. S. Winlaw, and R. P. Harvey.** 2007. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. Am J Hum Genet **81:**280-291.

25. **Kohlhase, J., M. Heinrich, L. Schubert, M. Liebers, A. Kispert, F. Laccone, P. Turnpenny, R. M. Winter, and W. Reardon.** 2002. Okihiro syndrome is caused by SALL4 mutations. Hum Mol Genet **11:**2979-2987.

26. **Kohlhase, J., L. Schubert, M. Liebers, A. Rauch, K. Becker, S. N. Mohammed, R. Newbury-Ecob, and W. Reardon.** 2003. Mutations at the SALL4 locus on chromosome 20 result in a range of clinically overlapping phenotypes, including Okihiro syndrome, Holt-Oram syndrome, acro-renal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. J Med Genet **40:**473-478.

27. **Kraus, F., B. Haenig, and A. Kispert.** 2001. Cloning and expression analysis of the mouse T-box gene *tbx20*. Mech Dev **100:**87-91.

28. **Kuhnlein, R. P., G. Frommer, M. Friedrich, M. Gonzalez-Gaitan, A. Weber, J. F. Wagner-Bernholz, W. J. Gehring, H. Jackle, and R. Schuh.** 1994. *spalt* encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. Embo J **13:**168-179.

29. **Luscher, B., P. J. Mitchell, T. Williams, and R. Tjian.** 1989. Regulation of transcription factor AP-2 by the morphogen retinoic acid and by second messengers. Genes Dev **3:**1507-1517.

30. **Lyons, I., et al.** 1995. Myogenic and morphogenic defects in the heart tubes of murine embryos lacking the homeobox gene *Nkx2.5*. Genes Dev. **9:**1654-1666.

31. **Mani, A., J. Radhakrishnan, A. Farhi, K. S. Carew, C. A. Warnes, C. Nelson-Williams, R. W. Day, B. Pober, M. W. State, and R. P. Lifton.** 2005. Syndromic patent ductus arteriosus: evidence for haploinsufficient TFAP2B mutations and identification of a linked sleep disorder. Proc Natl Acad Sci U S A **102:**2975-2979.

32. **McKinsey, T. A., and E. N. Olson.** 2005. Toward transcriptional therapies for the failing heart: chemical screens to modulate genes. J Clin Invest **115:**538-546.

33. **Meins, M., D. J. Henderson, S. S. Bhattacharya, and J. C. Sowden.** 2000. Characterization of the human TBX20 gene, a new member of the T-Box gene family closely related to the Drosophila H15 gene. Genomics **67:**317-332.

34. **Mori, A. D., and B. G. Bruneau.** 2004. TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol **19:**211-215.

35. **Moser, M., J. Ruschoff, and R. Buettner.** 1997. Comparative analysis of AP-2 alpha and AP-2 beta gene expression during murine embryogenesis. Dev Dyn **208:**115-124.

36. **Okubo, A., O. Miyoshi, K. Baba, M. Takagi, K. Tsukamoto, A. Kinoshita, K. Yoshiura, T. Kishino, T. Ohta, N. Niikawa, and N. Matsumoto.** 2004. A novel GATA4 mutation completely segregated with atrial septal defect in a large Japanese family. J Med Genet **41:**e97.

37. **Papaioannou, V. E., and L. M. Silver.** 1998. The T-box gene family. BioEssays **20:**9-19.

38. **Parmacek, M. S., and J. A. Epstein.** 2005. Pursuing cardiac progenitors: regeneration redux. Cell **120:**295-298.

39. **Pashmforoush, M., J. T. Lu, H. Chen, T. S. Amand, R. Kondo, S. Pradervand, S. M. Evans, B. Clark, J. R. Feramisco, W. Giles, S. Y. Ho, D. W. Benson, M. Silberbach, W. Shou, and K. R. Chien.** 2004. *Nkx2-5* pathways and congenital heart disease; loss of ventricular myocyte lineage specification leads to progressive cardiomyopathy and complete heart block. Cell **117:**373-386.

40. **Pehlivan, T., B. R. Pober, M. Brueckner, S. Garrett, R. Slaugh, R. Van Rheeden, D. B. Wilson, M. S. Watson, and A. V. Hing.** 1999. GATA4 haploinsufficiency in patients with interstitial deletion of chromosome region 8p23.1 and congenital heart disease. Am J Med Genet **83:**201-206.

41. **Richardson, P., W. McKenna, M. Bristow, B. Maisch, B. Mautner, J. O'Connell, E. Olsen, G. Thiene, J. Goodwin, I. Gyarfas, I. Martin, and P. Nordet.** 1996. Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. Circulation **93:**841-842.

42. **Satoda, M., F. Zhao, G. A. Diaz, J. Burn, J. Goodship, H. R. Davidson, M. E. Pierpont, and B. D. Gelb.** 2000. Mutations in TFAP2B cause Char syndrome, a familial form of patent ductus arteriosus. Nat Genet **25:**42-46.

43. **Schonberger, J., H. Levy, E. Grunig, S. Sangwatanaroj, D. Fatkin, C. MacRae, H. Stacker, C. Halpin, R. Eavey, E. F. Philbin, H. Katus, J. G. Seidman, and C. E. Seidman.** 2000. Dilated cardiomyopathy and sensorineural hearing loss: a heritable syndrome that maps to 6q23-24. Circulation **101:**1812-1818.

44. **Schott, J., D. W. Benson, C. T. Bassson, W. Pease, S. Silberbach, J. P. Moak, B. J. Maron, C. E. Seidman, and J. G. Seidman.** 1998. Congenital Heart Disease caused by mutations in the transcription factor NKX2-5. Science **281:**108-111.

45. **Singh, M. K., V. M. Christoffels, J. M. Dias, M. O. Trowe, M. Petry, K. Schuster-Gossler, A. Burger, J. Ericson, and A. Kispert.** 2005. Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. Development **132:**2697-2707.

46. **Stennard, F. A., M. W. Costa, D. A. Elliott, S. Rankin, S. J. Haast, D. Lai, L. P. McDonald, K. Niederreither, P. Dolle, B. G. Bruneau, A. M. Zorn, and R. P. Harvey.** 2003. Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart. Dev Biol **262:**206-224.

47. **Stennard, F. A., M. W. Costa, D. Lai, C. Biben, M. B. Furtado, M. J. Solloway, D. J. McCulley, C. Leimena, J. I. Preis, S. L. Dunwoodie, D. E. Elliott, O. W. Prall, B. L. Black, D. Fatkin, and R. P. Harvey.** 2005. Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. Development **132**:2451-2462.

48. **Takeuchi, J. K., M. Mileikovskaia, K. Koshiba-Takeuchi, A. B. Heidt, A. D. Mori, E. P. Arruda, M. Gertsenstein, R. Georges, L. Davidson, R. Mo, C. C. Hui, R. M. Henkelman, M. Nemer, B. L. Black, A. Nagy, and B. G. Bruneau.** 2005. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. Development **132**:2463-2474.

49. **Wayne, S., N. G. Robertson, F. DeClau, N. Chen, K. Verhoeven, S. Prasad, L. Tranebjarg, C. C. Morton, A. F. Ryan, G. Van Camp, and R. J. Smith.** 2001. Mutations in the transcriptional activator EYA4 cause late-onset deafness at the DFNA10 locus. Hum Mol Genet **10:**195-200.

50. **Yagi, H., Y. Furutani, H. Hamada, T. Sasaki, S. Asakawa, S. Minoshima, F. Ichida, K. Joo, M. Kimura, S. Imamura, N. Kamatani, K. Momma, A. Takao, M. Nakazawa, N. Shimizu, and R. Matsuoka.** 2003. Role of TBX1 in human del22q11.2 syndrome. Lancet **362:**1366-1373.

51. **Yamagishi, H., and D. Srivastava.** 2003. Unraveling the genetic and developmental mysteries of 22q11 deletion syndrome. Trends Mol Med **9:**383-389.

52. **Zhao, F., C. G. Weismann, M. Satoda, M. E. Pierpont, E. Sweeney, E. M. Thompson, and B. D. Gelb.** 2001. Novel TFAP2B mutations that cause Char syndrome provide a genotype-phenotype correlation. Am J Hum Genet **69:**695-703.

CHAPTER 3

DEVELOPMENTAL EXPRESSION PATTERNS OF *TBX1*, *TBX2*, *TBX5*, AND *TBX20* IN *XENOPUS TROPICALIS*

PREFACE

Chapter 3 focuses on the evolutionary conservation of T-box gene orthologs in *X. tropicalis*, and emphasizes its utility as a model organism for studies of gene function. While T-box genes such as *Tbx5* and *Tbx20* have been previously characterized in *X. laevis*, their expression pattern in the closely related diploid species *X. tropicalis* remained unexplored. Through these studies, the importance and utility of *X. tropicalis* in genetic studies of gene regulation is emphasized.

This work was previously produced and published in collaboration with Dr. Christopher Showell and Kathleen S. Christine. I was directly involved with the work pertaining to *Tbx20* sequence conservation and expression. Additionally, I participated in the writing and editing of the manuscript for publication.

Showell, C., K.S. Christine, E.M. Mandel, and F.L. Conlon. 2006. Developmental expression patterns of *Tbx1*, *Tbx2*, *Tbx5*, and *Tbx20* in *Xenopus tropicalis*. Developmental Dynamics **235(6)**:1623-1630.

SUMMARY

T-box genes have diverse functions during embryogenesis and are implicated in several human congenital disorders. Here we report the identification, sequence analysis and developmental expression patterns of four members of the T-box gene family in the diploid frog *Xenopus tropicalis*. These four genes – *Tbx1*, *Tbx2*, *Tbx5* and *Tbx20* – have been shown to influence cardiac development in a variety of organisms, in addition to their individual roles in regulating other aspects of embryonic development. Our results highlight the high degree of evolutionary conservation between orthologues of these genes in *X.tropicalis* and other vertebrates, both at the molecular level and in their developmental expression patterns, and also identify novel features of their expression. Thus, *X.tropicalis* represents a potentially valuable vertebrate model in which to further investigate the functions of these genes through genetic approaches.

INTRODUCTION

DNA-binding transcription factors encoded by several members of the T-box gene family have been shown to have both cell-autonomous and non-cell autonomous roles in controlling the development of the heart during embryogenesis (30, 39). These roles appear to be conserved during evolution and, in some cases, their importance is highlighted by the association between mutations in these factors and the incidence of human congenital heart defects (22, 27, 32). In addition, the same genes have been shown to be required for the proper development of other tissues and organs, such as the eye (*Tbx5*) (17) and ear (*Tbx1*) (20, 24, 28, 31, 41), while other T-box genes have key roles in regulating early embryonic patterning (35).

Xenopus is a valuable model organism in which to investigate the molecular and genetic regulation of organogenesis in general and heart development in particular, and reverse genetic approaches have recently been developed to isolate mutant alleles in specific genes of interest in the diploid frog *Xenopus (Silurana) tropicalis*. In comparison with the zebrafish (*Danio rerio*), *Xenopus* cardiac morphology is more similar to that of humans, including septation of the atrium into left and right chambers (12, 23). Also, the accessibility of the embryo throughout development and the high fecundity of the frog are significant advantages over the mouse, both in embryological analysis and in genetic screening.

The genes analyzed here – *Tbx1, Tbx2, Tbx5 and Tbx20* – are all known to play important roles in regulating normal cardiac development. *TBX1* lies within a critical region of human chromosome 22 (22q11.2) that is deleted in patients with DiGeorge syndrome, and loss of Tbx1 function in the mouse mimics the severe morphological defects of the outflow tract of the heart that are seen in DiGeorge patients (14, 21). Similarly, mutations in the human *TBX5* gene are associated with Holt-Oram syndrome, affecting atrioventricular septation, the cardiac conduction system and the development of the upper limbs (4, 19). *Tbx5* has been shown to act in concert with *Tbx20* at the molecular level to control cardiac morphogenesis (6). Conversely, *Tbx5* and *Tbx2* appear to function within distinct domains of the developing heart, contributing to the patterning of the early heart tube and its subsequent morphological regionalization. A number of recent studies have also demonstrated a requirement for *Tbx20* function for proper regulation of *Tbx2* expression within the developing heart (7, 36, 38). As a preliminary step in investigating the molecular basis of their developmental roles through genetic analysis in the emerging model organism *Xenopus tropicalis*, we have identified cDNA clones containing full-length coding sequences

corresponding to these four T-box genes, determined the structure of their genomic loci *in silico*, and characterized their spatial expression patterns over a wide range of stages during embryogenesis. Our results demonstrate the high degree of sequence conservation of T-box gene orthologues in *Xenopus tropicalis* and highlight both conserved and previously undescribed aspects of their embryonic expression.

RESULTS AND DISCUSSION

Sequence analysis of *Xenopus tropicalis* T-box gene orthologues

cDNA clones corresponding to *Tbx1*, *Tbx2* and *Tbx20* were identified by BLAST searches within a database of *Xenopus tropicalis* expressed sequence tags and a clone containing the translation initiation codon was obtained and sequenced for each gene. A cDNA encoding the *X.tropicalis* orthologue of Tbx5 was cloned by RT-PCR and sequenced. These cDNA sequences were used to search the *X.tropicalis* draft genome sequence (DoE Joint Genome Institute) for genomic scaffolds containing the corresponding loci. The cDNA sequences were then mapped onto the genomic locus sequences and the exon/intron boundaries were identified based on consensus sequences for eukaryotic splice donor and acceptor sites (25).

All four *Xenopus tropicalis* cDNA clones exhibit a very high degree of sequence identity when compared with their *Xenopus laevis* orthologues, particularly within their coding regions. The 1389nt open reading frame within the 3065bp *Tbx1* cDNA is 94% identical to that of *Xenopus laevis Tbx1* (Genbank Acc. # AF526274) (89% identity in untranslated regions). The degrees of identity and similarity between the conceptually translated *Xenopus tropicalis Tbx1* coding sequence and several vertebrate orthologues are

shown in Table 3.1. The results of our analysis of the genomic *Tbx1* locus are shown in Figure 3.1a.

The 3510bp *Tbx2* cDNA identified here contains a 2055nt open reading frame with 94% identity to *Xenopus laevis Tbx2* (Genbank Acc. # AB032941) (86% identity in untranslated regions). Table 3.1 shows the degrees of identity and similarity between conceptually translated *Xenopus tropicalis* Tbx2 and vertebrate orthologues. Mapping of the *Tbx2* cDNA sequence to the available genome sequence identified a 14,129bp region containing the complete cDNA sequence divided amongst seven exons (Fig. 3.1b).

Xenopus tropicalis Tbx5 is encoded by a 1557nt open reading frame. Alignment of this sequence with the *Xenopus laevis Tbx5* cDNA (Genbank Acc. # AF133036) identified 93% nucleotide sequence identity between the coding regions of the two orthologues. The *Xenopus tropicalis Tbx5* cDNA encodes a product exhibiting a high degree of evolutionary conservation amongst vertebrate species (Table 3.1). Results obtained from *in silico* analysis of the *Tbx5* genomic locus are shown in Figure 2.1c.

The *Tbx20* cDNA clone obtained consists of 2117bp, containing a 1320nt open reading frame with 93% sequence identity to that of *Xenopus laevis Tbx20* (Genbank Acc. # AY154394) (75% identity in untranslated regions). Table 3.1 shows the degree of sequence identity and similarity between conceptually translated *Xenopus tropicalis* Tbx20 and its orthologues in other vertebrates. The *Tbx20* sequence was found to be divided amongst eight exons within a 19,354bp region of a single genomic scaffold (Fig. 3.1d).

Analysis of *Tbx1* expression during embryogenesis

Tbx1 function is required for normal heart development in vertebrates. It is thought to act indirectly, influencing the differentiation of migrating cardiac neural crest cells by regulating the expression of one or more intercellular signals emanating from *Tbx1*-expressing cells in the pharyngeal endoderm and the mesenchymal core of the pharyngeal arches (16). The cardiac neural crest cells contribute to the formation of the outflow tract of the heart and the development of this region is severely affected in DiGeorge patients and in mouse models of the syndrome. Initial analysis of the phenotype of a hypomorphic $Tbx1^{neo}$ allele in the mouse suggests that the observed alignment and septation defects of the outflow tract are independent, thus underscoring the value of analyzing more subtle alleles in addition to single gene knockouts and larger deletions in vertebrate models (42). To determine the spatial patterns of *Tbx1* mRNA expression during the course of *X.tropicalis* embryogenesis, whole mount in situ hybridization was performed. At the earliest stage analyzed, stage 10.5 (early gastrula), no expression of *Tbx1* was detected. In early neurulae (stage 13), regionally restricted expression was clearly detected in a broad anterior domain surrounding the anterior end of the medio-dorsal groove of the neural plate (Fig. 3.2a). Within this broad ectodermal domain, two bilateral patches of strong *Tbx1* expression were detected flanking the medio-dorsal groove (Fig. 3.2a,c). These patches marked the posterior boundary of the *Tbx1* expression domain. In late neurulae (stage 19), strong expression was detected in the anterior ectoderm (Fig. 3.2d-f). This expression domain appeared to largely exclude the central nervous system, commonly defined by the expression of pan-neural markers such as the neural cell adhesion molecule (N-CAM) (8). Expression was not detected in the developing eye anlagen and cement gland, and was only weakly detected in the region of the neural tube

posterior to the eye anlagen. Instead, expression of *Tbx1* was found to immediately abut these regions of the ectoderm. As at stage 13, two distinct bilateral regions of strong staining were observed within the *Tbx1* expression domain at stage 19, extending as approximately dorsoventral stripes in the ectoderm on either side of the anterior CNS. It is unclear whether this *Tbx1* expression domain corresponds to the location of the proposed primordium of the ectodermal (neurogenic) placodes (33, 34). At early tailbud stage (stage 25), *Tbx1* was found to be expressed in three distinct areas within the pharyngeal region and in the ventral region of each otic vesicle (Fig. 3.2g,h). At stage 33, expression within the otic vesicles extended further laterally (Fig. 3.2n). However, in subsequent stages (stages 40, 47) expression remained restricted to the ventral and lateral regions of the vesicles. This differs from the pattern reported for *X.laevis*, in which *Tbx1* appeared to be expressed throughout the vesicles (3).

Expression of *Tbx1* orthologues in the pharyngeal region is broadly conserved amongst vertebrate species. Between stages 25 and 33, the elaboration of the expression pattern of *Tbx1* in this region of the *X.tropicalis* embryo reflects the morphogenesis of the pharyngeal arches. In this region, the cells expressing *Tbx1* lay beneath the overlying epidermis. At stages 25 and 26, expression was detected in the mandibular and hyoid arches (Fig. 3.2g,h,i) and in a third domain corresponding to the future branchial arches, posterior to the hyoid arch. At stage 27, at which the first branchial arch becomes fully formed, *Tbx1* expression was detected in four distinct pharyngeal domains – the mandibular, hyoid and first branchial arches and a more posterior branchial region (Fig. 3.2j,k). By stage 33, expression was also detected in the second branchial arch (Fig. 3.2n). At this stage, *Tbx1* appears to

mark distinct dorsal and ventral regions within the hyoid, first branchial, second branchial and forming third branchial arches.

Analysis of *Tbx2* expression during embryogenesis

In situ hybridization showed that, as in *X.laevis* (10), *Tbx2* is expressed ventrally in *X.tropicalis* early gastrulae (stage 10.5) (Fig. 3.3a,b). However, in contrast to the reported expression in *X.laevis*, *Tbx2* is expressed most strongly in the outer layer of ectodermal cells in *X.tropicalis* (Fig. 3.3c). In dissected wholemount embryos and in cryosectioned embryos (Fig. 3.3c), very faint staining was observed in the underlying ventral mesoderm. At the late gastrula stage (stage 12), expression appeared to be consistently upregulated in a small group of cells clustered around the ventral edge of the closing blastopore (Fig. 3.3d). At the early neurula stage (stage 13), four regions of ectodermal expression were clearly detected. Strong staining was observed in the developing cement gland (Fig. 3.3e) and in a U-shaped domain around the proctodeum at the posterior of the embryo (Fig. 3.3g). Two bilateral patches of expression were seen in the head, at the edge of the neural plate, in the region of the future neurogenic placodes caudal to the eye anlagen (Fig. 3.3f,g). It is unclear whether this domain includes both the profundal-trigeminal placodal area and the dorsolateral placodes. In *Xenopus*, the dorsolateral placodes give rise to the lateral line placodes and the otic placodes at later stages (34). Finally, a diffuse pattern of *Tbx2*-positive cells was seen in the ventral epidermis (Fig. 3.3h, j-l). At stage 19 (late neurula), expression persists in the cement gland, the ventral epidermis, the proctodeal region, the lens placodes, and in a broad placodal area caudal to the eye anlagen (Fig. 3.3h-l). In addition, expression was detected in the dorsal root ganglia of the future spinal cord (Fig. 3.3i). At stage 21/22, *Tbx2* expression was seen in a

wishbone-shaped group of cells situated dorsal and caudal to each developing optic vesicle (Fig. 3.3m), corresponding to the cranial (profundal and trigeminal) ganglia. Expression was found to persist in these cells through tailbud and into early tadpole stages (Fig. 3.3n-r). From stage 21/22 onwards, the bilateral expression of *Tbx2* in the ectodermal placodes became restricted primarily to the otic placode and the developing otic vesicles. Unlike *Tbx1*, *Tbx2* was found to be expressed throughout the otic vesicles, and this expression was detected at all subsequent stages analyzed (stages 24 to 40). At stage 24, additional staining was observed in the precursors of the hypaxial muscles and the pronephric duct in the trunk, in the developing branchial arches, and in the primordium of the heart (Fig. 3.3o,p). A small group of cells within the telencephalon is also stained at this stage (Fig. 3.3o-r). In stage 29 embryos, expression was clearly detected in the frontonasal process (Fig. 3.3q). *Tbx2* continues to be expressed in the same regions of the embryo at stage 33, although its expression becomes clearly regionalized in the looping heart. A higher level of expression was clearly detected in the ventricle compared with the atrium, as reported in other organisms (Fig. 3.3r).

Analysis of *Tbx5* expression during embryogenesis

The expression pattern of *Tbx5* was analyzed at developmental stages from mid-gastrula (stage 11) to early tadpole (stage 40). No expression was detected at stage 11. In late neurulae (stage 19), a gradient of *Tbx5* expression was present within the eye anlagen, with higher levels dorsally (Fig. 3.4a,b). At this stage, two small patches of cells on either side of the embryo were also stained, corresponding to regions within the migrating bilateral heart primordia. At early tailbud stage (stage 25), this pattern of expression was maintained in a

dorsal region of each developing eye (Fig. 3.4c) and in the heart primordia, located ventrally (Fig. 3.4c,d). In stage 26 embryos, the *Tbx5*-expressing cells of the heart primordia were seen to converge at the ventral midline (Fig. 3.4f), while expression was also detected in two bilateral groups of cells continuous with and extending dorsally from the heart primordia. These cells likely correspond to the progenitors of the right and left branches of the sinus venosus and common cardinal veins (11, 26). In other organisms, *Tbx5* has been shown to play an important role in eye development, particularly in guiding the projection of neurons between the retina and tectum (17). In *X.tropicalis*, expression in the dorsal region of the eye was found to be maintained until early tadpole stages (stage 40), although its expression becomes greatly restricted between stages 33 and 40 (Fig. 3.4h,i). At stage 31/32, strong expression was detected in the posterior region of the heart tube in cleared embryos (Fig. 3.4g). Following looping of the heart, a higher level of expression was detected in the ventricle (situated ventrally and offset to the left side of the embryo) than in the atrium (Fig. 3.4j). The regional differences in the expression of *Tbx5* within the hearts of *X.tropicalis* tadpoles were seen consistently in both wholemount and sectioned embryos. Transverse sections through the heart at stages following heart looping showed expression of *Tbx5* in the ventricular myocardium, while staining was not detected in the atrial region of the heart (Fig. 3.6b).

Analysis of *Tbx20* expression during embryogenesis

The expression pattern of *X.tropicalis Tbx20* was analyzed in embryos between stages 13 (neural plate stage) and 40 (early tadpole). Although expressed weakly in the developing cement gland as early as stage 13 in *X.laevis* (5), we did not detect expression at this stage in

X.tropicalis. At late neurula stage (stage 20), *Tbx20* was strongly expressed in the developing cement gland and in the bilateral heart primordia (Fig. 3.5a-c). Expression in the cement gland was found to decrease from stage 25 onwards, while expression continued to be strongly detected in the developing heart. In the heart-forming region at stage 25, a single domain of expression was detected, corresponding to the heart field formed by fusion of the bilateral heart primordial (Fig. 3.5e). This fusion of the *Tbx20*-expressing domains appears to occur earlier in *X.tropicalis* than in *X.laevis*. Notably, this pattern of expression differs considerably from that of *Tbx5*, in which fusion of the bilateral pre-cardiac expression domains begins at around stage 26 (see above). In addition to this cardiac expression, two small domains of expression were observed in the hindbrain (rhombencephalon) at this stage, corresponding to the second and fourth rhombomeres. At stage 29/30, expression persisted in these regions and was also weakly detected in a more posterior region of the hindbrain (Fig. 3.5f). The hindbrain expression of *Tbx20* was found to be upregulated in embryos at subsequent stages and, as in more anterior regions, was detected in distinct paired subdomains (Fig. 3.5g,j,k,m,n). At stage 33, when heart looping is initiated, *Tbx20* was found to be broadly expressed in the heart tube, with strong staining detected in the ventricle, atrium and both branches of the sinus venosus (inflow tract) (Fig. 3.5g,i). Thus, the expression domain of *Tbx20* in the developing chambers of the heart tube only partially overlaps that of *Tbx5*. This is consistent with the patterns of *Tbx20* expression reported in other vertebrates (1, 5, 18, 29, 43). During heart looping (stage 36), *Tbx20* was expressed at a higher level in the atrium than in the ventricle (Fig. 3.5l). This regional difference in the expression level of *Tbx20* was maintained in early tadpole stage embryos (stage 40) and was

clearly seen both in whole embryos and in transverse sections through the heart (Fig. 3.5o., 3.6c).

MATERIALS AND METHODS

Identification and isolation of cDNA clones

cDNA clones TNeu106g11, TGas050k23 and TTpA031n09, encoding *X.tropicalis* orthologues of *Tbx1*, *Tbx2* and *Tbx20* respectively, were identified by searching a database of *X.tropicalis* expressed sequence-tagged clones derived from oligo-dT primed cDNA libraries specific to several developmental stages (www.sanger.ac.uk; (9)). Specifically, nucleotide sequences from the 5' ends of the coding regions of the corresponding *X.laevis* orthologues (*Tbx1* Genbank Acc. # AF526274; *Tbx2* Genbank Acc. # AB023815; *Tbx20* Genbank Acc. # AY154394) were used to BLAST search (2) for *X.tropicalis* clones containing the predicted translation start codon and which were therefore likely to contain full-length cDNAs. The clones were obtained (MRC Geneservice) and the cDNA inserts were sequenced. A cDNA encoding Tbx5 was cloned by low-stringency RT-PCR, using total RNA template from stage 13-20 *X.tropicalis* embryos. Primers were designed based on sequences flanking the *X.laevis* *Tbx5* coding sequence (forward: 5'-GAAGATCTATGGCGGACACAGAGGAGGCT-3'; reverse: 5'-GAGAGATCTACGCTGTTTTCATTCCAGTCTGG-3'). The resulting product was cloned into *pcDNA3.1* (Invitrogen Corp.). All cDNA sequences are deposited in Genbank (*Tbx1* accession # DQ124205; *Tbx2* accession # DQ124206; *Tbx5* accession # DQ124207; *Tbx20* accession # DQ124208).

***In silico* analysis**

To identify genomic sequence scaffolds corresponding to *Tbx1*, *Tbx2*, *Tbx5* and *Tbx20*, the corresponding cDNA sequences were used to search the *X.tropicalis* draft genome sequence (versions 2.0 and 3.0) using the BLAST algorithm (2) (DoE Joint Genome Institute). Pairwise sequence alignments and analyses of sequence conservation of conceptually translated proteins were performed using GeneDoc (www.psc.edu/biomed/genedoc).

Embryo collection and *in situ* hybridization

X.tropicalis embryos were collected following natural single-pair mating between animals from a partially inbred (F6) line (NASCO). Males and females were pre-primed with ten units of human chorionic gonadotropin (hCG; Sigma) twenty hours before being primed with an additional two hundred units. One hour after priming, males and females were paired and allowed to mate for approximately five hours in shallow water at 25ºC. Embryos and unfertilized eggs from successful matings were collected, treated with 2% cysteine hydrochloride to remove their jelly coat, and sorted. Embryos were cultured at 25ºC in sterilized water from our aquatic system and staged according to criteria set out in the Normal Table of *Xenopus laevis* (26).

A 908bp *KpnI-XhoI* fragment of the *Tbx1* EST clone was subcloned into *pBluescript-KS* and this construct was linearized with *Acc657* to generate a template for in situ hybridization probe synthesis. Template for *Tbx2* probe synthesis was produced by linearizing the full-length cDNA clone described above using *HindIII*. The *Tbx5* cDNA was cut from *pcDNA3.1-Tbx5* by *NotI-SpeI* digest and sub-cloned into *pBluescript-KS* to generate

a probe template construct that was subsequently linearized with *NotI*. To generate a template for *Tbx20* probe synthesis, a 565bp *SalI-NotI* fragment from the *Tbx20* EST clone was sub-cloned into *pBluescript-KS* and the construct linearized using *SalI*. In situ hybridizations were performed according to a standard protocol (37) with the following exceptions: Fixed embryos were devitellinized by enzymatic treatment with collagenase A (Roche Applied Science), proteinase K and hyaluronidase (Sigma) (13). No further proteinase K treatment was performed. Embryos were pre-hybridized overnight (approx. 15 hours) and the RNase treatment step prior to antibody incubation was omitted (15). After staining with BM Purple alkaline phosphatase substrate (Roche Diagnostics), embryos were re-fixed in 1X MEM salts containing 10% formamide and then dehydrated in methanol.

Where necessary, embryos were cleared in 2:1 benzyl benzoate:benzyl alcohol (Sigma). Embryos were photographed on a Leica M-series stereomicroscope (Leica Microsystems Ltd.) using the Spot Advanced image capture system (Diagnostic Instruments Inc.) and edited using Photoshop 7.0 (Adobe Systems Inc.).

Cryosectioning

For cryosectioning, embryos were embedded in gelatin using a method modified from Stern and Holland (40). Following *in situ* hybridization, embryos were fixed in 4% paraformaldehyde in PBS and incubated overnight at 4ºC in 30% sucrose/PBS (w/v). The embryos were then pre-warmed to 38ºC before being transferred to 15% sucrose/PBS containing 7.5% gelatin (~300 Bloom; SIGMA) at 38ºC. Embryos were incubated in gelatin for a minimum of thirty minutes before being transferred to specimen molds (Tissue-Tek; Sakura Finetek U.S.A., Inc.). Embedded embryos were stored at 4ºC prior to cryosectioning.

Sections were taken at a thickness of 20μm. Gelatin was rinsed from the sections using PBS at 38ºC before mounting in aqueous mounting medium (Faramount; DakoCytomation).

ACKNOWLEDGEMENTS

EST clones were produced by Cambridge University, the Wellcome Trust Sanger Institute and Wellcome Trust/Cancer Research UK Institute *Xenopus tropicalis* EST Project. The authors would like to thank Dr. Victoria K. Graham (Duke University) and Sarah Goetz (UNC-Chapel Hill) for advice on cryosectioning, Dr. Eva Anton (UNC-Chapel Hill) for the use of equipment, and Misty Hurt (University of Virginia), Shruti Nagaraj and Shauna Vasilatos (UNC-Chapel Hill) for technical assistance. This work was supported by NIH/NHLBI grants HL075256-01 and HL083965-01 to F.L.C.

	Xenopus laevis	*Danio rerio*	*Gallus gallus*	*Mus musculus*	*Homo sapiens*
Tbx1	97% (98%)	80% (87%)	N/A	71% (79%)	69% (77%)
Tbx2	96% (97%)	77% (84%)	N/A	70% (77%)	70% (77%)
Tbx5	95% (96%)	64% (73%)	81% (87%)	78% (85%)	78% (85%)
Tbx20	97% (98%)	85% (92%)	90% (95%)	90% (95%)	N/A

Table 3.1. Sequence conservation of T-domain protein orthologues in vertebrates.

X.tropicalis Tbx1, *Tbx2*, *Tbx5* and *Tbx20* were analyzed by pairwise alignment with their orthologues in *X.laevis* (African clawed frog), *Danio rerio* (zebrafish), *Gallus gallus* (chicken), *Mus musculus* (mouse) and *Homo sapiens* (human). Overall sequence identity and similarity (in parentheses) between amino acid sequences are shown as percentages. Where putative full-length sequences were not available, these comparisons were omitted (N/A).

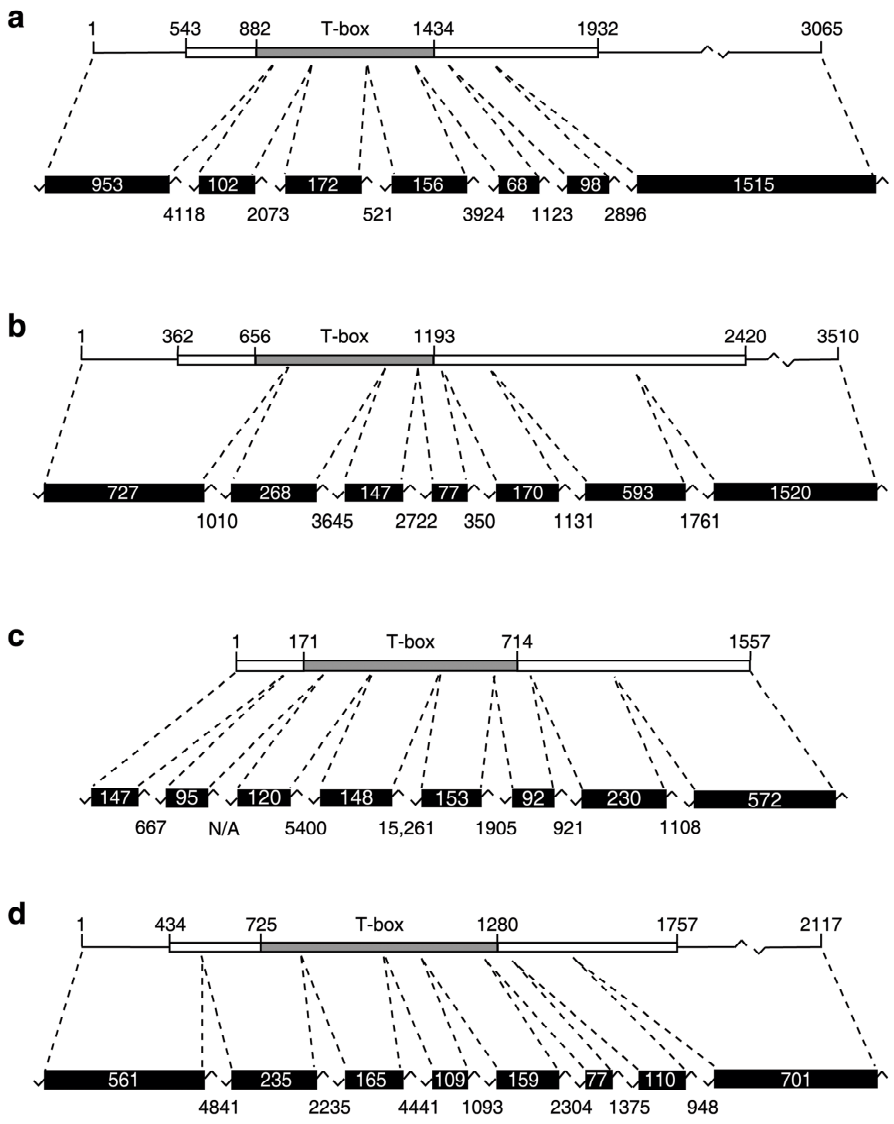


Figure 3.1. Genomic locus structure of *Tbx1*, *Tbx2*, *Tbx5* and *Tbx20* in *X.tropicalis*.

Tbx1, *Tbx2*, *Tbx5* and *Tbx20* cDNAs and their corresponding genomic loci are shown in diagrammatic form (not to scale). Coding regions of each cDNA are shown (boxes) together with their nucleotide positions and the position of the T-box (defined by alignment of the encoded proteins with the T-domain of Xbra) is also indicated. The exons corresponding to the cDNA sequences are shown together with their sizes (in base pairs) plus those of the intervening introns. Note that as the size of the first exon of each gene is predicted based on the available cDNA sequence, the sizes of these exons may be underestimated here.

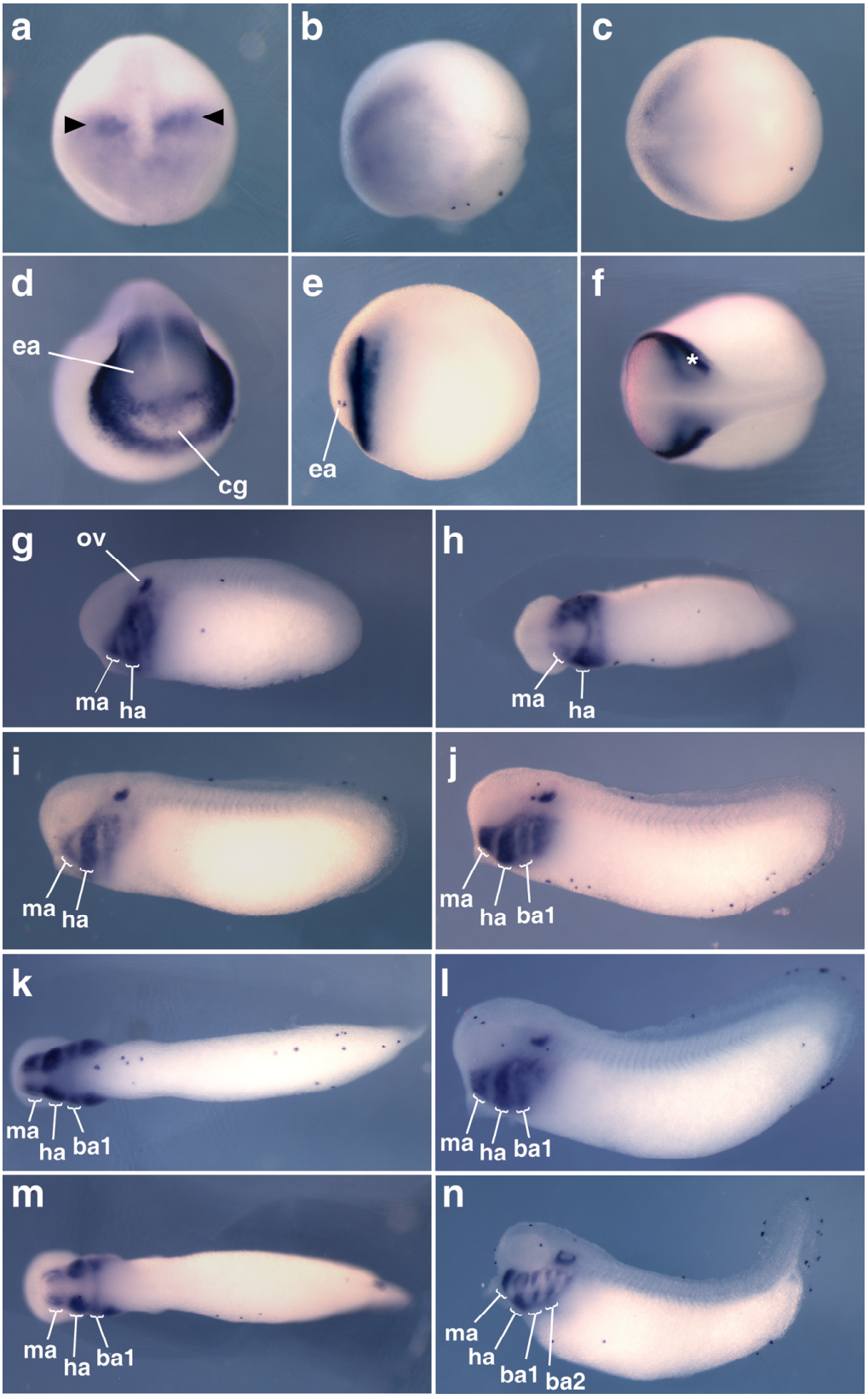

a
b
c
d
ea
cg
e
ea
f
*
g
ov
ma
ha
h
ma
ha
i
ma
ha
j
ma
ha
ba1
k
ma
ha
ba1
l
ma
ha
ba1
m
ma
ha
ba1
n
ma
ha
ba1
ba2

Figure 3.2. Expression pattern of *Tbx1* in *X.tropicalis*. The results of in situ hybridizations for *Tbx1* expression from early neurula to late tailbud stages are shown (embryos uncleared). Except for the anterior views shown in a) and d), all embryos are oriented with anterior to the left. Stage 13 is shown in anterior (a), lateral (b) and dorsal (c) views. Bilateral patches of stronger expression are indicated in a) by arrowheads. Stage 19 is shown in anterior (d), lateral (e) and dorsal (f) views. Bilateral stripes of stronger expression are indicated in f) by an asterisk. *Tbx1* expression through tailbud stages is shown as follows: Stage 25 lateral (g) and ventral (h), stage 26 lateral (i), stage 27 lateral (j) and ventral (k), stage 28 lateral (l) and ventral (m), stage 33 lateral (n). *ba1* first branchial arch, *ba2* second branchial arch, *cg* cement gland, *ea* eye anlagen, *ha* hyoid arch, *ma* mandibular arch, *ov* otic vesicle.

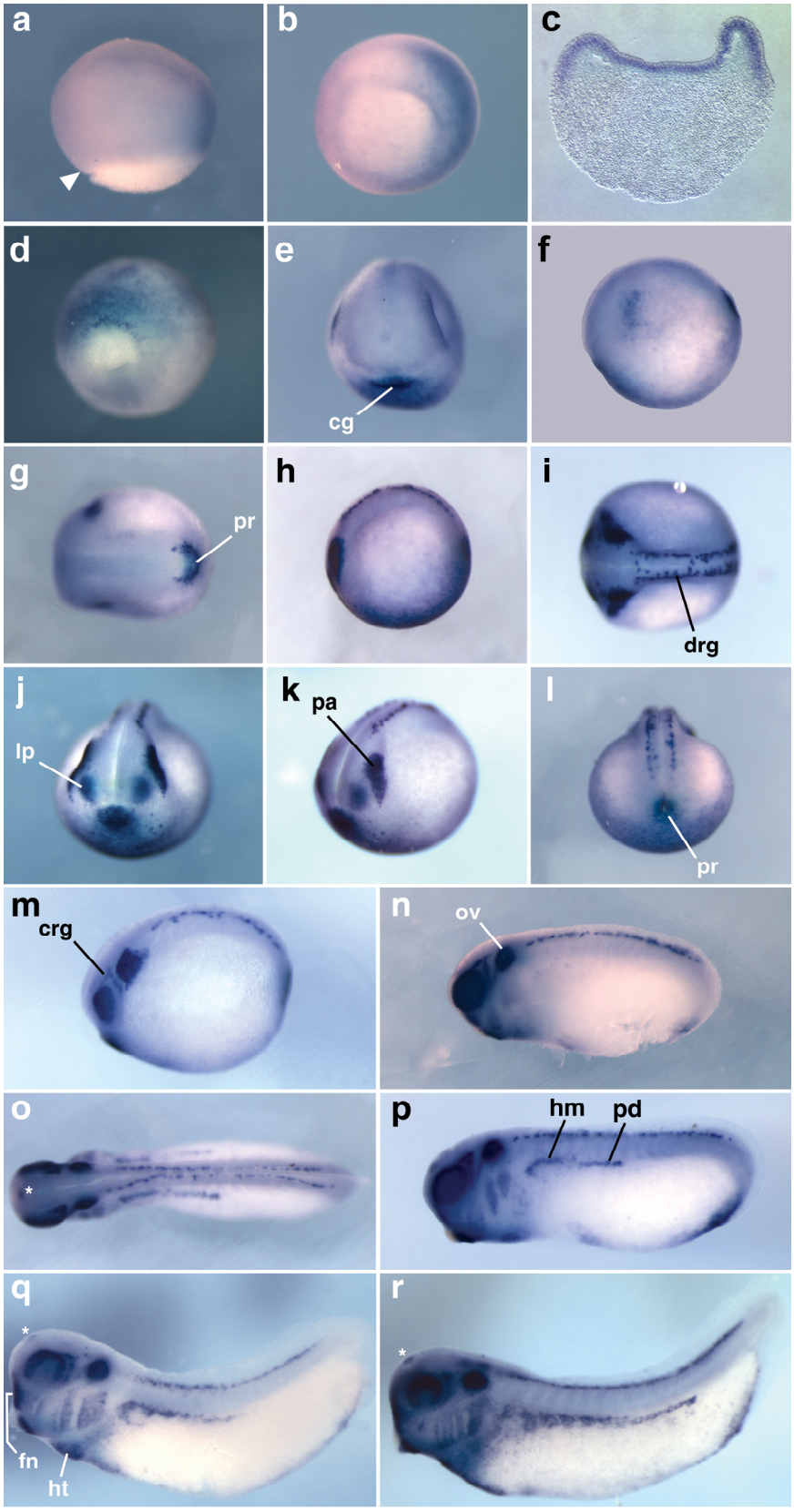

a
b
c
d
e
cg
f
g
pr
h
i
drg
j
lp
k
pa
l
pr
m
crg
n
ov
o
*
p
hm
pd
q
*
fn
ht
r
*

Figure 3.3. Expression pattern of *Tbx2* in *X.tropicalis*. *In-situ* hybridization results are shown for *Tbx2* (embryos uncleared). Expression at early gastrula (stage 10.5) is shown in lateral (a) and vegetal (b) views of wholemount embryos, and in transverse section (c; ventral to the right). In both a) and b), the embryo is oriented with dorsal to the left and the dorsal blastopore lip is indicated by an arrowhead in a). A vegetal view of a late gastrula (st12) is shown in d), ventral side uppermost. Expression at early neurula (stage 13) (e-g) late neurula (stage 19) (h,l) and tailbud stages 21/22 (m), 25 (n), 26 (o,p), 29 (q) and 33 (r) are also shown. Expression in the forebrain (telencephalon) at tailbud stages is indicated by an asterisk (o-r). Except for anterior (d,j) and posterior (l) views, all embryos are oriented with anterior to the left. *cg* cement gland, *crg* cranial ganglia, *drg* dorsal root ganglia, *fn* frontonasal process, *hm* hypaxial muscle, *ht* heart tube, *lp* lens placode, *ov* otic vesicle, *pd* pronephric duct, *pr* proctodeum.

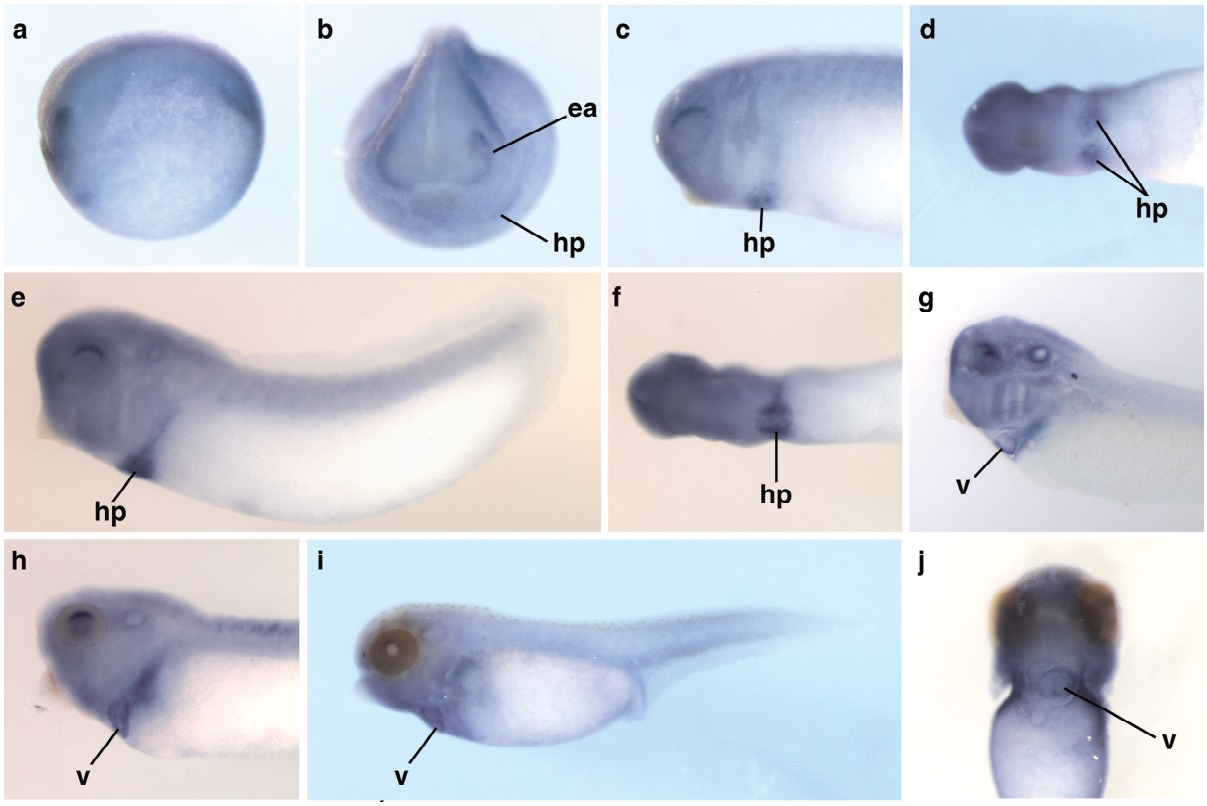


Figure 3.4. Expression pattern of *Tbx5* in *X.tropicalis*. The expression pattern of *Tbx5* detected by in situ hybridization between late neurula and early tadpole stages is shown (uncleared except for g). Stages are as follows: Stage 19 (a,b), stage 25 (c,d), stage 26 (e,f), stage 31/32 (g) (cleared), stage 33 (h), and stage 40 (i,j). Except for the anterior view in b) and the ventral view in j), embryos are oriented with anterior to the left. Anterior is to the top in j). *ea* eye anlagen, *hp* heart primordium, *v* ventricle.

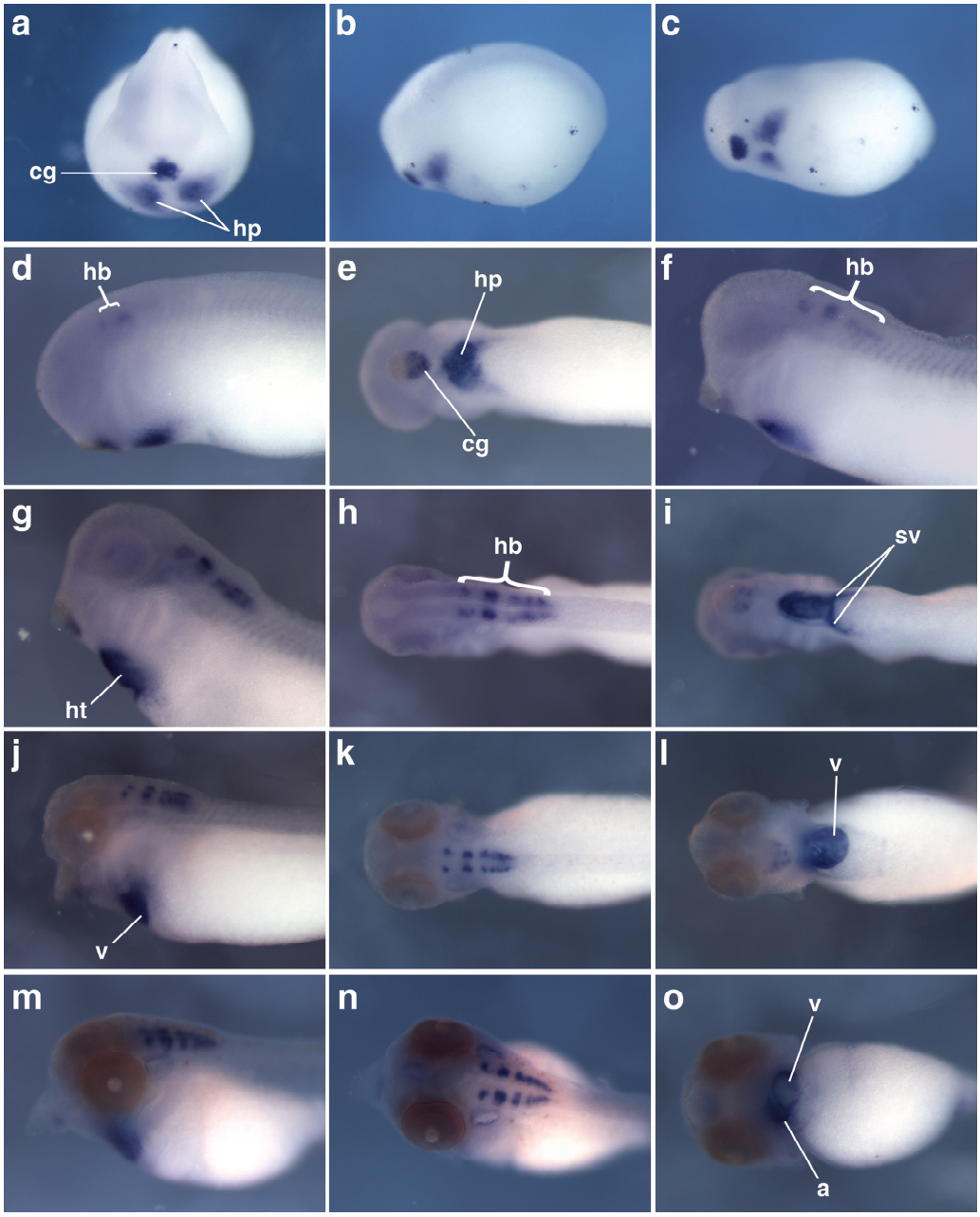


Figure 3.5. Expression pattern of *Tbx20* in *X.tropicalis*. The expression pattern of *Tbx20* detected by in situ hybridization between late neurula and early tadpole stages is shown (embryos uncleared). Stages are as follows: Stage 20 (a-c), stage 25 (d, e), stage 29/30 (f), stage 33 (g-i), stage 36 (j-l), and stage 40 (m-o). Except for the anterior view in a), embryos are oriented with anterior to the left. *a* atrium, *cg* cement gland, *h* heart, *hb* hindbrain, *hp* heart primordium, *ht* heart tube, *sv* sinus venosus, *v* ventricle.

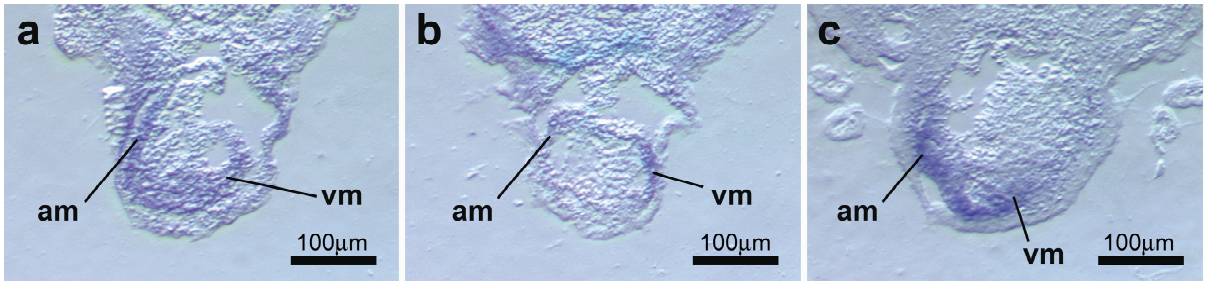


Figure 3.6. Cardiac expression of *Tbx2, Tbx5* and *Tbx20*. The *in situ* hybridization patterns of *Tbx2*, *Tbx5* and *Tbx20* in the forming cardiac chambers were examined in transverse sections through the tadpole heart after looping. a) *Tbx2* expression was seen in the myocardium of both the atrial (*am*) and ventricular (*vm*) regions of the looped heart at stage 36. b) Expression of *Tbx5* was restricted primarily to the developing ventricular myocardium at stage 38. c) In contrast to *Tbx5*, high levels of *Tbx20* expression were seen in the atrial region but not in the ventricle (stage 40). Magnification: 100X.

REFERENCES

1. **Ahn, D. G., I. Ruvinsky, A. C. Oates, L. M. Silver, and R. K. Ho.** 2000. *tbx20*, a new vertebrate T-box gene expressed in the cranial motor neurons and developing cardiovascular structures in zebrafish. Mech Dev **95:**253-258.

2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J Mol Biol **215:**403-410.

3. **Ataliotis, P., S. Ivins, T. J. Mohun, and P. J. Scambler.** 2005. XTbx1 is a transcriptional activator involved in head and pharyngeal arch development in *Xenopus laevis*. Dev Dyn **232:**979-991.

4. **Basson, C. T., D. R. Bachinsky, R. C. Lin, T. Levi, J. A. Elkins, J. Soults, D. Grayzel, E. Kroumpouzou, T. A. Traill, J. Leblanc-Straceski, B. Renault, R. Kucherlapati, J. G. Seidman, and C. E. Seidman.** 1997. Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. Nat Genet **15:**30-35.

5. **Brown, D. D., O. Binder, M. Pagratis, B. A. Parr, and F. L. Conlon.** 2003. Developmental expression of the *Xenopus laevis Tbx20* orthologue. Dev Genes Evol **212:**604-607.

6. **Brown, D. D., S. N. Martz, O. Binder, S. C. Goetz, B. M. J. Price, J. C. Smith, and F. L. Conlon.** 2005. *Tbx5* and *Tbx20* act synergistically to control vertebrate heart morphogenesis. Development **132**:553-563.

7. **Cai, C. L., W. Zhou, L. Yang, L. Bu, Y. Qyang, X. Zhang, X. Li, M. G. Rosenfeld, J. Chen, and S. Evans.** 2005. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development **132:**2475-2487.

8. **Eagleson, G., B. Ferreiro, and W. A. Harris.** 1995. Fate of the Anterior Neural Ridge and the Morphogenesis of the *Xenopus* Forebrain. Journal of Neurobiology **28:**146-158.

9. **Gilchrist, M. J., A. M. Zorn, J. Voigt, J. C. Smith, N. Papalopulu, and E. Amaya.** 2004. Defining a large set of full-length clones from a *Xenopus tropicalis* EST project. Developmental Biology **271**:498-516.

10. **Hayata T, K. H., Eisaki A, Asashima M.** 1999. Expression of *Xenopus* T-box transcription factor, *Tbx2* in *Xenopus* embryo. Development Genes and Evolution **209**:625-628.

11. **Horb, M. E., and G. H. Thomsen.** 1999. *Tbx5* is essential for heart development. Development **126**:1739-1751.

12. **Hu, N., D. Sedmera, H. J. Yost, and E. B. Clark.** 2000. Structure and function of the developing zebrafish heart. Anat Rec **260:**148-157.

13. **Islam, N., Moss, T.** 1996. Enzymatic removal of vitelline membrane and other protocol modifications for whole mount *in situ* hybridization of *Xenopus* embryos. Trends in Genetics **12:**459.

14. **Jerome, L. A., and V. E. Pappaioannou.** 2001. DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. Nature Genetics **27:**286-291.

15. **Khokha, M. K., C. Chung, E. L. Bustamante, L. W. Gaw, K. A. Trott, J. Yeh, N. Lim, J. C. Lin, N. Taverner, E. Amaya, N. Papalopulu, J. C. Smith, A. M. Zorn, R. M. Harland, and T. C. Grammer.** 2002. Techniques and probes for the study of *Xenopus tropicalis* development. Dev Dyn **225:**499-510.

16. **Kochilas, L., J. Liao, S. Merscher-Gomez, R. Kucherlapati, B. Morrow, and J. A. Epstein.** 2005. New insights into the role of *Tbx1* in the DiGeorge mouse model. *In* M. Artman, D. W. Benson, D. Srivastava, and M. Nakazawa (ed.), Cardiovascular Development and Congenital Malformations: Molecular and Genetic Mechanisms. Blackwell Publishing.

17. **Koshiba-Takeuchi, K., J. K. Takeuchi, K. Matsumoto, T. Momose, K. Uno, V. Hoepker, K. Ogura, N. Takahashi, H. Nakamura, K. Yasuda, and T. Ogura.** 2000. Tbx5 and the retinotectum projection. Science **287**:134-137.

18. **Kraus, F., B. Haenig, and A. Kispert.** 2001. Cloning and expression analysis of the mouse T-box gene *tbx20*. Mech Dev **100:**87-91.

19. **Li, Q. Y., R. A. Newbury-Ecob, J. A. Terrett, D. I. Wilson, A. R. Curtis, C. H. Yi, T. Gebuhr, P. J. Bullen, S. C. Robson, T. Strachan, D. Bonnet, S. Lyonnet, I. D. Young, J. A. Raeburn, A. J. Buckler, D. J. Law, and J. D. Brook.** 1997. Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. Nat Genet **15:**21-29.

20. **Liao, J., L. Kochilas, S. Nowotschin, J. S. Arnold, V. S. Aggarwal, J. A. Epstein, M. C. Brown, J. Adams, and B. E. Morrow.** 2004. Full spectrum of malformations in velo-cardio-facial syndrome/DiGeorge syndrome mouse models by altering *Tbx1* dosage. Hum Mol Genet **13:**1577-1585.

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

22. **Mandel, E. M., T. E. Callis, D.-Z. Wang, and F. L. Conlon.** 2005. Transcriptional mechanisms of congenital heart disease. Drug Discovery Today: Disease Mechanisms **2(1)**:33-38.

23. **Mohun, T. J., L. M. Leong, W. J. Weninger, and D. B. Sparrow.** 2000. The morphology of heart development in *Xenopus laevis*. Dev Biol **218:**74-88.

24. **Moraes, F., A. Novoa, L. A. Jerome-Majewska, V. E. Papaioannou, and M. Mallo.** 2005. Tbx1 is required for proper neural crest migration and to stabilize spatial patterns during middle and inner ear development. Mech Dev **122:**199-212.

25. **Mount, S.** 1982. A catalogue of splice junction sequences. Nucleic Acids Research **10:**459-472.

26. **Nieuwkoop, P. D., and J. Faber (ed.).** 1967. Normal Table of *Xenopus laevis (Daudin)*, 2nd ed. North-Holland Pub. Co., Amsterdam.

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

28. **Piotrowski, T., D. G. Ahn, T. F. Schilling, S. Nair, I. Ruvinsky, R. Geisler, G. J. Rauch, P. Haffter, L. I. Zon, Y. Zhou, H. Foott, I. B. Dawid, and R. K. Ho.** 2003. The zebrafish *van gogh* mutation disrupts *tbx1*, which is involved in the DiGeorge deletion syndrome in humans. Development **130:**5043-5052.

29. **Plageman, T. F., Jr., and K. E. Yutzey.** 2004. Differential expression and function of Tbx5 and Tbx20 in cardiac development. J Biol Chem **279:**19026-34.

30. **Plageman, T. F., Jr., and K. E. Yutzey.** 2005. T-box genes and heart development: putting the "T" in heart. Dev Dyn **232:**11-20.

31. **Raft, S., S. Nowotschin, J. Liao, and B. E. Morrow.** 2004. Suppression of neural fate and control of inner ear morphogenesis by Tbx1. Development **131:**1801-1812.

32. **Ryan, K., and A. J. Chin.** 2003. T-box genes and cardiac development. Birth Defects Res C Embryo Today **69:**25-37.

33. **Schlosser, G., and K. Ahrens.** 2004. Molecular anatomy of placode development in *Xenopus laevis*. Developmental Biology **271**:439-466.

34. **Schlosser, G., and R. G. Northcutt.** 2000. Development of Neurogenic Placodes in *Xenopus laevis*. Journal of Comparative Neurology **418:**121-146.

35. **Showell, C., O. Binder, and F. L. Conlon.** 2004. T-box genes in early embryogenesis. Dev Dyn **229:**201-18.

36. **Singh, M. K., V. M. Christoffels, J. M. Dias, M. O. Trowe, M. Petry, K. Schuster-Gossler, A. Burger, J. Ericson, and A. Kispert.** 2005. Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. Development **132:**2697-2707.

37. **Sive, H., Grainger RM, Harland RM.** 2000. Early Development of *Xenopus laevis* - A Laboratory Manual. Cold Spring Harbor Laboratory Press.

38. **Stennard, F. A., M. W. Costa, D. Lai, C. Biben, M. B. Furtado, M. J. Solloway, D. J. McCulley, C. Leimena, J. I. Preis, S. L. Dunwoodie, D. E. Elliott, O. W. Prall, B. L. Black, D. Fatkin, and R. P. Harvey.** 2005. Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. Development **132:**2451-2462.

39. **Stennard, F. A., and R. P. Harvey.** 2005. T-box transcription factors and their roles in regulatory hierarchies in the developing heart. Development **132:**4897-4910.

40. **Stern, C. D., and P. W. H. Holland (ed.).** 1993. Essential Developmental Biology: A Practical Approach. IRL Press at Oxford University Press.

41. **Vitelli, F., A. Viola, M. Morishima, T. Pramparo, A. Baldini, and E. Lindsay.** 2003. TBX1 is required for inner ear morphogenesis. Hum Mol Genet **12:**2041-2048.

42. **Xu, H., M. Morishima, and A. Baldini.** 2005. *Tbx1* and DiGeorge syndrome: a genetic link between cardiovascular and pharyngeal development, Cardiovascular Development and Congenital Malformations: Molecular and Genetic Mechanisms. Blackwell Publishing.

43. **Yamagishi, T., Y. Nakajima, S. Nishimatsu, T. Nohno, K. Ando, and H. Nakamura.** 2004. Expression of *tbx20* RNA during chick heart development. Dev Dyn **230:**576-80.

CHAPTER 4

EXPRESSION VIA A NOVEL REGULATORY ELEMENT OF THE CARDIAC GENE *Tbx20* IS REGULATED BY BMP SIGNALING

PREFACE

Chapter 4 describes the growth factor controlled regulation of *Tbx20* expression in the *Xenopus* heart. With *Nkx2.5* and *Tbx5*, *Tbx20* is one of the earliest markers expressed in the developing heart and has been shown to be essential for proper cardiac development. However, until now, the signaling pathways involved in the tissue-specific expression of *Tbx20* remained unknown. Here, we demonstrate that a conserved, 334bp regulatory element located directly upstream of *Tbx20* is both necessary and sufficient to drive its expression in the heart. Additionally, we show that this element is regulated by BMP signaling through direct binding of SMAD1 to multiple SMAD binding sites. Thus, this work represents the first characterization of *Tbx20* regulation in the heart, and further expands our understanding of the transcriptional regulatory networks that govern cardiogenesis.

This work represents a collaborative effort with Thomas E. Callis, Dr. Da-Zhi Wang, Natalie Thomas and Dr. Deborah Yelon and was recently submitted for publication to the journal *Circulation Research*. For this project, I performed the BAC library screen, immunohistochemistry, reporter construct cloning, *Xenopus* transgenesis, DNA binding (Pherastar) assays and completed the writing of the manuscript.

Mandel, E.M., T.E. Callis, N. Thomas, D. Yelon, D.Z. Wang and F.L. Conlon. 2007. Expression via a novel regulatory element of the cardiac gene *Tbx20* is regulated by BMP signaling. Circ. Res. In Review.

SUMMARY

The T-box gene, *Tbx20* has been demonstrated to be mutated in human congenital heart disease and has been shown to be required for early heart development in a wide array of model systems including zebrafish, *Xenopus* and mouse. Despite the evolutionarily conserved requirement for TBX20 in the control of a range of cardiac cellular processes, the signal transduction pathways that lie upstream of *Tbx20* remain completely unknown. Here, we describe the first example of specific growth factor-controlled regulation of *Tbx20* gene expression in the developing heart. We have identified and characterized a conserved 334bp regulatory element located upstream of the *Tbx20* locus that is both necessary and sufficient to drive cardiac specific expression of *Tbx20* in *Xenopus* and zebrafish. We further demonstrate that activation of *Tbx20* is dependent on a novel regulatory element that contains seven distinct canonical and non-canonical SMAD high affinity binding sites. This therefore represents evidence for a direct role of the BMP/SMAD signaling pathway in the regulation of cardiac *Tbx20* expression.

INTRODUCTION

Members of the T-box protein family have been demonstrated to play critical roles in heart development and disease. For example, patients with Holt-Oram Syndrome (HOS)

syndrome frequently lack *Tbx5* (15, 34). Recently, clinical studies have also implicated the T-box containing protein TBX20 in human congenital heart disease (21). Specifically, familial mis-sense or non-sense mutations within the DNA binding domain or T-box of *Tbx20* are associated with a wide range of cardiac defects in humans including septal defects, abnormal valve formation and altered cardiomyocyte growth (21).

Studies of *Tbx20* orthologues in a wide range of model systems including mouse (*Tbx12/20*) (7, 22), zebrafish (*Tbx20/HrT*) (1, 14), chick (19), and *Xenopus* (3) have shown that along with *Tbx5* and *Nkx2.5*, *Tbx20* is one of the earliest genes expressed in the vertebrate cardiac lineage. *Tbx20* continues to be expressed in the heart until adulthood and moreover, appears to be expressed in all regions of the developing heart including the atria, ventricles, inflow and outflow tracts, and the septum transversum. Consistent with these findings, as well as clinical studies of human congenital heart disease, *Tbx20* orthologues have been shown to be required for proper cardiogenesis in zebrafish (49), *Xenopus* (4), and mouse (5, 44, 47, 50). Thus, the sequence, expression and function of *Tbx20* appear to be evolutionarily conserved.

Despite the critical role for *Tbx20* in heart development and disease, very little is known about the signal transduction pathways that function upstream to regulate *Tbx20* expression in cardiac tissue. However, studies on the regulation of T-box gene during gastrulation have suggested that many of the signal transduction pathways that are necessary for early cardiac induction, such as transforming growth factor-β (TGFβ) and Wnt, may also be function to activate or maintain T-box expression. For example, *Brachyury*, the founding member of the T-box gene family is expressed throughout newly formed mesoderm, with *Xenopus Brachyury* (*Xbra*) expression induced in tissue explants in an immediate early

response to the TGF-β family member activin (45). Subsequent studies have gone on to show that TGF-β signaling is required for *Xbra* expression *in vivo* (17). Based on these findings, the region of the *Xbra* promoter was cloned and shown to contain elements that confer a dose-dependent response to activin (9, 26, 27, 48).

In addition to activin, members the TGF-β family include nodal, anti-Muellerian hormone (AMH) and bone morphogenetic proteins (BMPs), all of which function by binding to specific heteromeric serine/threonine kinase receptor complexes located at the cell membrane. This association leads to the phosphorylation of SMAD molecules, the cytoplasmic mediators of the signal transduction pathway. Specifically, TGF-β, nodal and activin molecules induce the activation of the receptor-regulated SMADs 2 and 3, while alternatively, the AMH and BMP molecules signal through receptor-regulated SMADs 1, 5 and 8. Once phosphorylated, these receptor-regulated SMADs associate with the common mediator SMAD, SMAD4, which facilitates translocation to the nucleus. Once in the nucleus, they bind target DNA sequences and, in conjunction with transcriptional co-factors, regulate the transcription of downstream target genes (reviewed in (23, 33)).

There is substantial evidence for a role for BMP molecules in early heart development. Initial evidence comes from the observation that mutation of the *Drosophila* BMP ortholog Dpp, or mutation in any component of the Dpp signaling cascade, leads to a loss of cardiac progenitor cells and thus, a failure of the fly embryo to develop a dorsal vessel, the structure analogous to the vertebrate heart (46). Consistent with these findings, promoter analysis of the gene *tinman*, a gene expressed in the dorsal vessel and required for its formation, has shown that induction of *tinman* expression requires Dpp, and its effectors Mad (SMAD1) and Medea (SMAD4), which directly bind to elements within the *tinman*

enhancer (55). The role of BMPs in vertebrate cardiac development is also supported by the phenotypic analysis of BMP pathway mutations in the mouse. For example, conditional mutations in the BMP receptor ALK3 in cardiomyocytes lead to a thinning of the trabeculae and abnormal cushion formation (12). The function of these molecules is further emphasized by the observation that mutations in genes within the BMP pathway are associated with human congenital heart disease (20, 56). However, due to potential redundancy within the BMP family, and the early lethality of embryos mutant for a number of BMP signaling pathway components, the precise role for any one member of the BMP family in cardiac development remains unclear (40).

To characterize the growth factor pathways that lead to cardiac expression of *Tbx20*, we have begun to define the cardiac regulatory elements within the *Tbx20* locus. This work has led to the identification of a 334bp regulatory element that is both necessary and sufficient for *Tbx20* expression in the heart during cardiac chamber formation. We go on to demonstrate that this element is sufficient to drive cardiac-specific expression in living zebrafish, demonstrating that the signal transduction pathways acting upstream of *Tbx20* are evolutionarily conserved between fish and frogs. We further show that this minimal cardiac element is regulated by the BMP pathway and its mediators SMAD1 and SMAD4, but not by the TGF-β/Activin/Nodal mediator SMAD3, and that *Tbx20* is co-expressed with nuclear localized SMAD1 during cardiac chamber formation. Finally, we demonstrate that the minimal cardiac *Tbx20* element contains seven canonical and non-canonical high affinity SMAD binding sites that are directly bound by SMAD1 protein and are necessary in a combinatorial fashion for proper regulation of *Tbx20*. Collectively, these studies support a role for BMP molecules such as BMP2/4 during cardiac chamber formation and provide the

first insights into the pathways by which cardiac signaling cascades regulate *Tbx20* expression.

MATERIALS AND METHODS

In Situ Hybridization and Immunohistochemistry

X. laevis embryos were collected and fixed in 4% parformaldehyde in phosphate buffered saline (PBS) for 2hrs. at room temperature and stored at 4^{o}C in PBS for up to one week. Embryos were placed in a 30% sucrose/PBS solution at 4^{o}C overnight before being embedded and frozen in OCT cryosectioning medium (TissueTek). *In situ* hybridization was performed on 20μm transverse cryostat sections using a DIG-labeled *Tbx20* RNA probe (3). The RNA probe was detected using an anti-digoxigenin Fab-AP antibody (Roche, 1:2000) following the manufacturer's protocol. For immunohistochemical analysis, frozen transverse sections (14μm) were rinsed with wash buffer (1% Triton X-100, 1% heat-inactivated lamb serum in PBS) and incubated overnight at 4^{o}C with one of the following primary antibodies diluted in wash buffer: rabbit anti-GFP (AbCam, 1:5000), mouse anti-tropomyosin (Developmental Studies Hybridoma Bank, 1:50), or rabbit anti-Smad 1/5/8 (Cell Signaling, 1:50). The sections were then rinsed with wash buffer and incubated for 30mins. at room temperature with the corresponding secondary antibody diluted in wash buffer as follows: Alexa 488 anti-rabbit (Molecular Probes, 1:500) or Cy3 anti-rabbit (Sigma, 1:100). Sections were then rinsed with wash buffer and incubated for 30mins. at room temperature with DAPI (Sigma). Histological sections were viewed using a Nikon Eclipse E800 microscope.

BAC Library Screen, RLM-RACE

The ISB-1 *X. tropicalis* bacterial artificial chromosome (BAC) library [Children's Hospital Oakland Research Institute (CHORI)] was screened using a 317bp PCR product amplified from the 5' terminus of the *X. laevis Tbx20* coding region, and BAC DNA prepared according to protocols described by CHORI. DNA was digested with a series of restriction enzymes, and analyzed using field inversion gel electrophoresis (FIGE) and Southern blot analysis with a *Tbx20* probe. From this analysis, a 4114bp EcoRI fragment of *Tbx20* was identified and ligated into the pBSII-KS+ vector (Stratagene). The transcriptional start site of *X. tropicalis Tbx20* was identified using the First Choice RLM-RACE Kit (Ambion) and 5' RLM-RACE as described by the manufacturer using whole *X. tropicalis* embryos (N=25) as well as brain-enriched and heart-enriched tissues (approximately 250 embryos for both) at stage 28. RNA was isolated using Trizol Reagent (Invitrogen).

***Tbx20*-EGFP Reporter Constructs**

Tbx20-EGFP reporter constructs for transgenesis were generated by introducing a BamHI site into Exon 1 of *Tbx20* at position +142 in the Tbx20-pBSII-KS+ construct using the QuickChange Site-Directed Mutagenesis Kit (Stratagene). Following BamHI digestion and re-ligation, an AflII site was introduced into the pBSII-KS+ backbone of the modified construct upstream of the NotI site. The EGFP reporter contained within the pEGFP-N1 vector (Clontech) was then isolated by BamHI/AflII digestion and subcloned into the modified *Tbx20*-pBSII-KS+ construct by BamHI/AflII to create the *Tbx20*(-2464)-EGFP reporter vector containing 2601bp of the 5' end of *Tbx20* driving an EGFP reporter (GenBank Accession #1047274). A *Tbx20*-EGFP deletion series was generated by

substituting elements of *Tbx20* ranging from 471-2106bp, each containing a 5' EcoRI linker and a 3' BamHI linker, for the original 2601bp of the *Tbx20*-EGFP construct. Details and primer sequences are available upon request.

Xenopus laevis Embryo Culture and *Xenopus* Transgenesis

Wildtype *Xenopus laevis* embryos were obtained and cultured using standard methods. Staging was performed according to Nieuwkoop and Faber (35). The *Tbx20* reporter constructs described were linearized by KpnI, and transgenesis performed according to previously described methods for restriction enzyme-mediated integration (REMI) (24). EGFP expression was analyzed in living embryos using a Leica MZFLIII fluorescent dissecting microscope. For each transgene described, greater than 10 EGFP positive embryos were examined in each of at least three independent sets of injections.

Zebrafish Injections

For injection of zebrafish transgenes, reporter constructs flanked by I-SceI meganuclease recognition sites were created from the *Tbx20*-EGFP reporter constructs utilized for *X. laevis* transgenesis (13). Details and primers are available upon request. *Tbx20*-EGFP constructs were digested with I-SceI, diluted in dH_2O and phenol red in 0.2M KCl to concentrations between 0.063 ng/nl and 0.24 ng/nl. Between 0.5 nl and 1.5 nl of each construct was injected into the cytoplasm of single cell zebrafish embryos. Injected embryos were screened at 48hpf for GFP expression on a Zeiss M2Bio microscope. Injections with each construct were repeated 4-6 times and data from embryos screened for GFP expression were pooled for each.

Cell Culture and Luciferase Assays

Tbx20-luciferase reporter constructs were generated by cloning the respective regions from the EGFP reporter constructs into pGL3-Basic (Invitrogen). For transfection, COS7 cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin, and transfected using Fugene6 (Roche) reagent. Cells were transfected with 100ng reporter plasmid and indicated expression plasmids. Expression plasmids used for the transcriptional assays have been previously described as follows:

Myocardin (51, 52)

SRF (51, 52)

pGL3-Nkx2.5 (30)

Mef2c (53)

Gata4 (36)

pRK5 N-Flag Smad1 (31)

Smad4 (11)

Smad3 (11)

SM22 (28)

Additionally, the mouse *Tbx5* coding sequence was cloned into the pcDNA3.1-N-myc vector. A CMV-lacZ reporter was used as an internal control to normalize for transfection efficiencies, and total amount of DNA per well was kept constant by adding the corresponding amount of empty expression vector. Reporter assays were conducted in triplicate at least two times in 12-well plates. Fold induction was calculated as induction

compared to that of reporter alone, and error bars refer to the standard deviation of fold induction.

Protein-DNA Binding Assays

Protein Expression

Human Smad1 cDNA was subcloned from the pRK5 N-Flag SMAD1 construct into the pGEX-4T expression vector for the production of a glutathione *S*-transferase (GST) fusion protein, and overexpressed in *E. coli* (BL21) at 30^{o} C(31). The GST-SMAD1 fusion protein was induced with 100μM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18^{o} C overnight and purified using Glutathione Sepharose 4B (GE Heathcare). GST-SMAD1 was eluted (50mM Tris-HCl, pH 8.0, 10mM reduced glutathione) and dialyzed (50mM Tris-HCl, pH 8.0, 100mM NaCl) before concentration.

5'-FAM Oligos

For 2X coverage of the *Tbx20*(-334) regulatory element, 21 double-stranded, 30 basepair, 5'-FAM oligos were designed to overlap by 15 bases beginning from base -1 (Figure 4.6A). Additionally, *XVent* and SRF binding site oligos were designed as positive and negative controls respectively, based on previously published work (8, 18). Forward (5'-FAM) oligos and reverse strand oligos were resuspended (10 mM Tris, pH8, 150 mM NaCl and 1 mM EDTA) and annealed to prepare duplex DNA for binding assays.

Fluorescence Polarization

All fluorescence polarization experiments were performed in a PHERAstar microplate reader (BMG Labtechnologies, Durham, NC). Reactions were performed in triplicate in black, flat-bottom, half-area, 96-well plates at 25^{o}C. The 50μL reactions contained 250nmol/L 5'-FAM oligo and increasing concentrations of GST-SMAD1 (0-7416.67nmol/L) in reaction buffer (10mM Tris-HCl, pH 8.0, 100mM NaCl). Anisotropy was measured by excitation with vertically polarized light, using 490nm excitation and 520nm emission filters with the gain optimized for maximum signal and normalized to "no protein" controls. Data analysis was performed using SigmaPlot 8.0 software, and dissociation constants (K_d) determined for each oligo using the single rectangular I, 3 parameter equation $y=y_o + ax/(b+x)$ where b is equal to K_d.

RESULTS

A *Tbx20*-EGFP transgene recapitulates the endogenous expression of *Tbx20*

To determine the regulatory elements required for *Tbx20* cardiac expression, we isolated a putative *Tbx20* promoter region flanking the 5' end of the *Xenopus tropicalis* (*X. tropicalis*) *Tbx20* gene by screening the C.H.O.R.I. *X. tropicalis* ISB-1 Bacterial Artificial Chromosome (BAC) library using a 5' *Tbx20* specific probe. Field inversion gel electrophoresis (FIGE) and Southern blotting were used to isolate a 4116bp region of the *Tbx20* locus. 5' RLM-RACE on whole embryos (stage 28) and embryonic hearts (stage 28) identified a transcriptional start site at a position 287bp upstream of the translational start site, and therefore contained in our cloned region of *Tbx20* (Figure 4.1A).

To determine the ability of this genomic element to direct cardiac expression of *Tbx20*, we inserted an EGFP reporter cassette in-frame with the TBX20 translational start site (Figure 4.1A, B). Based on our observations that *Tbx20* is expressed in an identical pattern in *X. tropicalis* and *X. laevis* embryos in the cement gland, rhombomeres and heart from early stages of development (3, 43), we introduced the *Tbx20*(-2464)-EGFP reporter into *X. laevis* embryos by restriction enzyme-mediated integration (REMI). Consistent with this endogenous pattern of *X. tropicalis* and *X. laevis Tbx20* expression at these stages, the *Tbx20* reporter directed expression of EGFP to the developing heart and cement gland, (≥5 rounds of injections; n≥20 EGFP-expressing embryos per experiment)(Figure 4.1C-F). Specifically, EGFP expression was first observed in the cement gland at stage 24 and in the heart at stage 32 (data not shown). As the cells of the cement gland began to undergo apoptosis, expression of EGFP in the cement gland was gradually reduced to the point that it was absent by late tadpole stages (stages 47-48). In contrast, EGFP expression in the heart was maintained through chamber differentiation and heart looping (>stage 46; Figure 4.1C-F). EGFP expression, however, was never observed prior to early tadpole stages (stage 32). We also note that we never observe expression of the *Tbx20* reporter in other tissue types that express *Tbx20* including the hindbrain and eye. Thus, similar to the regulation of other cardiac genes such as *Nkx2.5* and *Mef2c*, multiple regulatory elements located throughout the *Tbx20* locus appear to be required for the complete temporal and spatial pattern of *Tbx20* expression.

334bp of upstream sequence is sufficient for *Tbx20*-EGFP expression

To begin to define the minimal element necessary for *Tbx20* cardiac expression in late tadpole stage embryos, we constructed a series of truncated reporters (Figure 4.2A). Examination of EGFP expression driven by these reporters reveals that 334bp of upstream sequence is sufficient to drive cardiac expression of the *Tbx20*-EGFP (≥3 rounds of injections; ≥20 EGFP-expressing embryos per experiment)(Figure 4.2B-J). We note that the same region is also sufficient to drive expression from stage 24 in the cement gland (Figure 4.2B-C, E-F, H-I).

To examine the precise tissue distribution of EGFP reporter expression, we serial sectioned and immunostained transgenic embryos. Each truncated reporter construct, (*Tbx20*(-2464)-EGFP, *Tbx20*(-1483)-EGFP and *Tbx20*(-334)-EGFP), recapitulated endogenous *Tbx20* expression throughout all cardiac tissue (8/8 EGFP-expressing embryos per construct)(Figure 4.2K-P). In addition, these studies reveal an increase in low levels of ubiquitous transgene expression throughout embryos carrying the 334bp reporter, suggesting the potential loss of negative regulatory elements located between -1483 and -334bp. Collectively, these data show that regions 5' to the *Tbx20* coding sequence are directly involved in the positive and negative regulation of *Tbx20* expression during early heart development. Moreover, only 334bp of upstream sequence is sufficient to drive the correct spatial cardiac-specific expression of *Tbx20*.

Signaling pathways upstream of the *Tbx20* cardiac-specific enhancer are evolutionarily conserved

To determine if the signal transduction pathways which function upstream of the *Tbx20* 334bp regulatory element are evolutionarily conserved, we analyzed the ability of *X. tropicalis Tbx20* regulatory elements to direct cardiac expression in zebrafish. We generated *Tbx20*(-2464) and *Tbx20*(-334)-EGFP fusion constructs flanked by I-SceI meganuclease sites (Figure 4.3A) and injected these transgenes together with the I-SceI meganuclease. Injection of reporter constructs in this fashion yields relatively efficient, yet highly mosaic, transgene expression (13). Injection of the construct containing the *Tbx20*(-2464) element into zebrafish embryos at the one cell stage resulted in EGFP expression in 68% (129/190) of embryos analyzed at 48 hours post fertilization (hpf) with 43% of these embryos expressing the *Tbx20* transgene in the heart (≥ 2 cells/heart, 56/129) (Figure 4.3C-D). Expression elsewhere in the embryos was inconsistent and seen in areas including the yolk, head and trunk. Injection of the smaller *Tbx20*(-334)-EGFP transgene also resulted in EGFP expression in the heart (Figure 4.3E-F). Of embryos analyzed at 48 hpf, 74% (251/338) of the embryos expressed the transgene, while 18% (45/251) of these embryos expressed EGFP only in the heart (Figure 4.3E-F). Again, expression outside of the heart was inconsistent. Therefore, these results show that the signaling pathway(s) utilized to drive cardiac expression of *Tbx20* in vertebrates are evolutionarily conserved.

***Tbx20* expression is regulated by SMAD1/SMAD4 but not SMAD3**

To begin to determine the pathways which lead to the cardiac expression of *Tbx20*, we examined the 334bp minimal *Tbx20* element by ConSite, Jaspar and Transfac software.

This analysis revealed putative transcription factor binding sites for a set of cardiac-specific transcription factors including SRF, NKX2.5, MEF2C, GATA4, SMAD1 and TBX5. To determine if any of these transcription factors could potentially regulate *Tbx20 in vivo*, we assayed their ability to activate transcription from cardiac-specific *Tbx20* reporters. To this end, *Tbx20*(-2464), *Tbx20*(-1483) and *Tbx20*(-334) luciferase reporter constructs were generated and tested for activation in transient transfections (Figure 4.4A-C). Of the potential transcription factors, only SMAD1, SRF, and the SRF co-factor myocardin, were capable of activating the *Tbx20* reporters (Figure 4.4A-C). Moreover, we observe that the 334bp reporter has a greater response to SMAD1 and SRF than the 1483bp or 2464bp elements suggesting sequences upstream of the 334bp element can attenuate the response to SMAD1 and SRF.

To more fully characterize the ability of the different members of the TGF-β family of signaling molecules to activate *Tbx20*, we tested the ability of SMAD1, a mediator of BMP signaling, SMAD3, a mediator of TGF-β/activin/nodal signaling, and the common SMAD, SMAD4, to activate *Tbx20* reporters. We observed a dose-dependent activation of both the largest (-2464) and smallest (-334) *Tbx20* regulatory elements when co-transfected with increasing amounts of SMAD1 (Figure 4.4D, G). High doses of SMAD1 induced a greater than 5-fold activation of the reporter, suggesting that the *Tbx20* regulatory element is responsive to the general BMP signaling pathway. Similarly, the *Tbx20*(-2464) and *Tbx20*(-334) cardiac regulatory elements are activated by SMAD4 in a dose-dependent manner (Figure 4.4E, H). In contrast, SMAD3 fails to activate any of the *Tbx20* reporter constructs while it is still able to induce expression of the SMAD3 target SM22 (Figure 4.4F)(38). To confirm that TGF-β/Activin/Nodal signaling is not involved in *Tbx20* induction, we also

utilized the small molecule inhibitor SB431542 (Sigma) to block signaling through the SMAD3-mediated activin pathway. In agreement with the lack of *Tbx20* induction by SMAD3, treatment with SB431542 did not reduce the induction of the *Tbx20* luciferase reporter by SMAD4, suggesting that *Tbx20* activation occurs in a SMAD1-dependent, SMAD3-independent manner (Figure 4.4I). Upon further analysis, we observed that *Tbx20* reporters did not respond to myocardin in a dose specific fashion *in vitro*, nor did that mutation of a putative SRF binding site have any effect on *Tbx20* reporter expression *in vivo*. Therefore, SRF and its co-factor myocardin were not analyzed in any further detail (data not shown).

Tbx20 and SMAD1 co-localize during cardiac chamber formation

Upon activation of the BMP signaling pathway SMAD1 is phosphorylated and translocates to the nucleus (reviewed in (23, 33)). To determine if SMAD1 is nuclear localized and co-expressed with *Tbx20*, and therefore could function endogenously to regulate *Tbx20* expression, we determined the precise spatial and temporal expression of SMAD1 relative to *Tbx20* during cardiac chamber formation. For these studies, serial sections of *X. laevis* hearts (stage 46) were examined by sectioned *in situ* hybridization for *Tbx20* expression and the adjacent sections stained for SMAD1. Results from these studies show that *Tbx20* is expressed throughout all cardiac tissue at stages of development corresponding to chamber formation (stage 46). Specifically, we observe *Tbx20* expression in the anterior regions of the heart, including the truncus arteriosus and atria, in a pattern that overlaps markers of the myocardium, such as tropomyosin (Figure 4.5A, C, E). *Tbx20* expression is also seen in mid and posterior portions of the heart in tissues including the

ventricle and outflow tract (Figure 4.5B, D, F). We observe an identical pattern of nuclear SMAD1 in adjacent sections. Thus, *Tbx20* and SMAD1 appear to be co-localized throughout all cardiac tissue types along the anterior-posterior and dorsal-ventral axes of the heart.

SMAD activation of *Tbx20* occurs through direct binding of SMAD1

SMAD proteins share two conserved Mad homology 1 (MH1) and Mad homology 2 (MH2) domains, of which the MH1 domain has been shown to bind DNA in a sequence-specific manner (16, 25). To identify the regions within the *Tbx20* cardiac regulatory element that are directly bound by SMAD1, we conducted fluorescence polarization DNA-binding assays. This technique can also be used to accurately determine the binding affinity of SMAD1 to the respective DNA sequence. To this end, we tiled double stranded, 30 bp, 5' carboxyfluorescein-labeled oligos covering the 334bp cardiac regulatory element overlapping by 15bp for full 2x coverage (Figure 4.6A). Based on the premise that oligos tumble more slowly in solution when bound by protein as compared to unbound oligos, we combined fluorescent oligos with increasing concentrations of a full length GST-SMAD1 fusion protein to evaluate the changes in light depolarization, as anisotropy, using a PHERAstar microplate reader (BMG Labtechnologies, Durham, NC). From the anisotropy data, we plotted binding curves and calculated dissociation constants (K_d) for each oligo interaction with SMAD1 (Figure 4.6B). As controls, we examined the binding of full length GST-SMAD1 to oligos containing known SMAD1 binding sequences (*XVent*) or known SRF binding sequences (8, 18). From these studies, seven individual oligos were bound more strongly than the *XVent* control oligo (Kd= 7.83μM) based on two standard deviations and a 95.4% confidence level. We note that dissociation constants calculated in these studies vary from those previously

shown for protein-DNA interactions, however, this is likely a reflection of the salt concentration required for the GST-SMAD1 protein to remain in solution (32, 54). SMAD1 bound oligo 8, covering bases -105 to -135, with the highest affinity (K_d=2.08μM). However, SMAD1 also bound six additional oligos (2, 6, 9, 13, 16, 19) with similar affinities (K_d=2.40μM to K_d=3.76μM)(Figure 4.6B). Results from these studies show that SMAD1 binds the *Tbx20* cardiac regulatory element at 7 individual sites (Figure 4.6C-D) and suggest that these seven sites are high affinity binding sites. This therefore supports a biologically relevant interaction between SMAD1 and sequences contained in the *Tbx20* reporter element (Figure 4.6B).

Sequence analysis of the oligos bound by SMAD1 reveals two conserved, consensus SMAD binding sites with the sequences GTCT and CAGAC in oligos 16 and 8, respectively. We observe that one region containing a putative binding site failed to bind SMAD1 protein *in vitro*. From this, we propose the presence of non-traditional SMAD binding elements between bases -15 and -45, -75 and -105, -120 and -150, -180 and -210, and -270 and -300. We further note that SMAD1 is capable of binding non-canonical sites with affinities equal to that of canonical sites.

Canonical SMAD sites alone are not sufficient for SMAD1 response

To determine the *in vivo* relevance of the canonical and non-canonical SMAD1 binding sites we first mutated the two conserved SMAD1 canonical sites either individually, or in combination, in the context of both the 334bp and 2464bp regulatory elements, yet saw little effect on *Tbx20* regulation in transcriptional assays or *X. laevis* transgenics (Figure 4.7A-B, Data not shown). To confirm and extend these findings, we generated additional

deletions of the 334bp element to 251bp. Thus, this reporter construct contains both canonical SMAD sites and all but one of the non-canonical sites. Similar to the *Tbx20*(-2464) or *Tbx20*(-334) regions, 251bp of *Tbx20* upstream sequence is responsive to SMAD1/4 with the *Tbx20*(-251) reporter induced greater than 20-fold by 200ng SMAD4 (Figure 4.7C). We note that these inductions are higher than observed for *Tbx20*(-334) reporter and most likely result from the removal of a negative regulatory element between -334bp and -251bp. When introduced into *X. laevis* the *Tbx20*(-251) reporter is expressed in the heart and cement gland in a pattern similar to that of the *Tbx20*(-334) reporter, however we also note an increase in EGFP expression in non-cardiac tissues (Figure 4.7E-F). A further deletion of the all canonical and all but one non-canonical SMAD1/4 binding sites greatly diminishes SMAD1/4 responsiveness and results in non-specific expression of the reporter in transgenic *X. laevis* (Figure 4.7D, G-H). Finally, if we delete the 334bp minimal element in the context of the original *Tbx20*(-2464) cardiac reporter, both the response to SMAD1/4 in tissue culture assays and EGFP expression in transgenic *X. laevis* embryos are abolished (Figure 4.7I, Data not shown).

Together, these data suggest that the 334bp region directly upstream of the *Tbx20* start site is both necessary and sufficient for cardiac expression of *Tbx20* and that *Tbx20* cardiac expression is dependent on SMAD1/4 activity. These data further suggest that at least one SMAD1/4 response element is located between -81 and -1bp upstream of the *Tbx20* transcriptional start site and additional elements are located between -251 and -81bp upstream of the start site, strongly suggesting that *Tbx20* expression in response to SMAD1/4 requires a complement of canonical and non-canonical SMAD1/4 binding sites.

DISCUSSION

Studies of cardiac gene regulation have suggested that heart specific transcription is regulated temporally and spatially through a set of distinct modular cis-acting elements (42). Here we report the identification of a 334bp element from *Xenopus* that contains a series of 7 high affinity SMAD1/4 binding sites that are necessary and sufficient for the evolutionarily conserved cardiac expression of *Tbx20*. Complementary to this finding, we have identified additional sequences that negatively attenuate the BMP/SMAD1 response. Taken together, these data strongly imply that BMP/SMAD1 signaling is necessary to maintain *Tbx20* expression during cardiac chamber formation.

Tbx20 cardiac expression and canonical and non-canonical SMAD1 binding sites

Our data demonstrate a requirement for a set of high affinity canonical and non-canonical SMAD binding sites in the regulation of *Tbx20* expression. Sequence analysis of the 334bp *Tbx20* cardiac element reveals two conserved, consensus SMAD binding sites with the sequences GTCT and CAGAC. We have demonstrated that mutation of the SMAD1/4 canonical binding sites either alone or in combination has little effect on the expression of *Tbx20* either *in vitro* or *in vivo*, implying that it is the complement of canonical and non-canonical SMAD1 binding sites which are required for the cardiac expression of *Tbx20*. Additionally, our results imply that the ability of SMAD1 to bind to DNA is not based on sequence alone. This is further supported by our observation that a region of the *Tbx20* minimal element containing a SMAD1/4 putative binding site failed to bind to SMAD1 protein *in vitro*. Taken together with our data in zebrafish, our data support a model in which evolutionarily conserved gene regulation by BMP involves a combinatorial set of unique

SMAD binding elements of which individual elements differ in their contribution to the response to growth factor signaling, and therefore transcriptional output.

Our findings that a complement of SMAD1/4 binding sties are required for *Tbx20* cardiac expression is broadly consistent with studies on two other BMP-responsive genes, *XVent* and *Nkx2.5*. In the case of *XVent*, the early mesodermal expression is dependent on five putative SMAD1/4 binding sites while cardiac expression of *Nkx2.5* is dependent on 12 individual SMAD1/4 binding sites (2, 18, 29, 30). Similar to our findings, point mutations or deletions of multiple SMAD1/4 binding sties in the *XVent* promoter has no effect on SMAD responsiveness (18). However, the cardiac-specific expression of *Tbx20* in cardiac tissue is unique in that, in the case of *Nkx2.5*, the cardiac specificity is mediated by a direct interaction between a SMAD1/SMAD4 complex and a member of the GATA transcription factor family (2). Although we have identified a GATA consensus site within the minimal *Tbx20* cardiac element, none of the *Tbx20* reporters respond to GATA4 (Figure 4.4), and deletion of the GATA site has no effect on cardiac-specific expression (Figure 4.7D). Thus, *Tbx20* cardiac-specific expression through the 334bp regulatory element, unlike that of *Nkx2.5*, appears to occur through a GATA-independent mechanism. The activation of cardiac gene expression via BMP signaling has also been shown to be dependent on additional cardiac transcription factors. For example, the myocardin-dependent expression of cardiac genes is synergistically activated by the direct interaction of SMAD1 with myocardin (6). However, our data with a large panel of cardiac transcription factors, demonstrate that with the exception of SRF, none of these factors could significantly induce *Tbx20* expression in transient transcriptional assays. We cannot formally rule out a potential role for SRF in *Tbx20* expression,

particularly in regulating the levels of *Tbx20*; however, mutation of the SRF site had no effect on *Tbx20* reporter expression *in vivo* (data not shown).

Cardiac-specific *Tbx20* expression

What then regulates the cardiac specific expression of *Tbx20*? We note that SMAD1 nuclear localization during heart development is temporally regulated. Based on this observation and our reporter analysis, we favor a model by which the complement of SMAD1/4 sites direct a broader pattern of expression in the embryo that is then restricted to the developing cardiac tissue by as of yet unidentified transcriptional repressors. This idea is supported by our observations that the SMAD1 and SMAD4 response of *Tbx20* reporters is enhanced upon deletion of regions both outside and within the 334bp element and that, upon reduction of the 334bp *Tbx20* regulatory element to 81bp, we observed a substantial increase in reporter expression in *X. laevis* transgenics in non-cardiac tissues. These findings are consistent with studies that have demonstrated that the BMP arm of the SMAD signaling pathway is associated with the regulation of genes involved in early heart development, while the TGF-β/activin/nodal arm of the SMAD pathway signals appear to drive cardiac regulation of factors associated with fibrotic, apoptotic and anti-hypertrophic events related to progression to heart failure (reviewed in (10)). Additionally, BMP signals have been suggested to act as long-range diffusible morphogens originating from multiple locations in the embryo including the endoderm, ectoderm or cardiac cells themselves (37, 39, 41). It is therefore interesting to speculate that the regulation of the novel 334bp *Tbx20* cardiac element during late cardiogenesis is a result of a continued or second wave of BMP signaling

from the underlying endoderm, or from the myocardial cells themselves mediated by SMAD1/4.

Although the SMAD1/4 sites are critical for the expression of *Tbx20* during cardiac chamber formation, the element of *Tbx20* that we have identified does not activate *Tbx20* in other regions of the embryo where *Tbx20* is endogenously expressed. Thus, our minimal *Tbx20* element does not comprise all of the appropriate information for the complete expression pattern of *Tbx20* and the early cardiac and neural elements of *Tbx20* are yet to be identified. From this, and the modularity of the BMP/SMAD response elements that we have characterized, it appears that like *Nkx2.5*, the regulation of *Tbx20* occurs in a modular manner. Finally, given that the minimal element that we have identified is required for expression of *Tbx20* during cardiac chamber formation, and given the correlation between mutations in *Tbx20* and human congenital heart disease (21), it will be interesting to determine if there exists an association between mutations in the *Tbx20* minimal element and congenital heart disease and/or cardiac hypertrophy.

ACKNOWLEDGEMENTS

We thank Dr. Shoko Ishibashi, Dr. Enrique Amaya, Scott A. Lujan, Dr. Laura M. Guogas and Dr. Matthew R. Redinbo for technical assistance, and Dr. Christopher Showell for critical reading of the manuscript and helpful suggestions. We would also like to thank Dr. Yonqin Wu and the UNC *In Situ* Hybridization Core Facility for valuable assistance. The tropomyosin antibody developed by J.-C. Lin was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by the University of Iowa, Department of Biological Sciences, Iowa City, IA 52242.

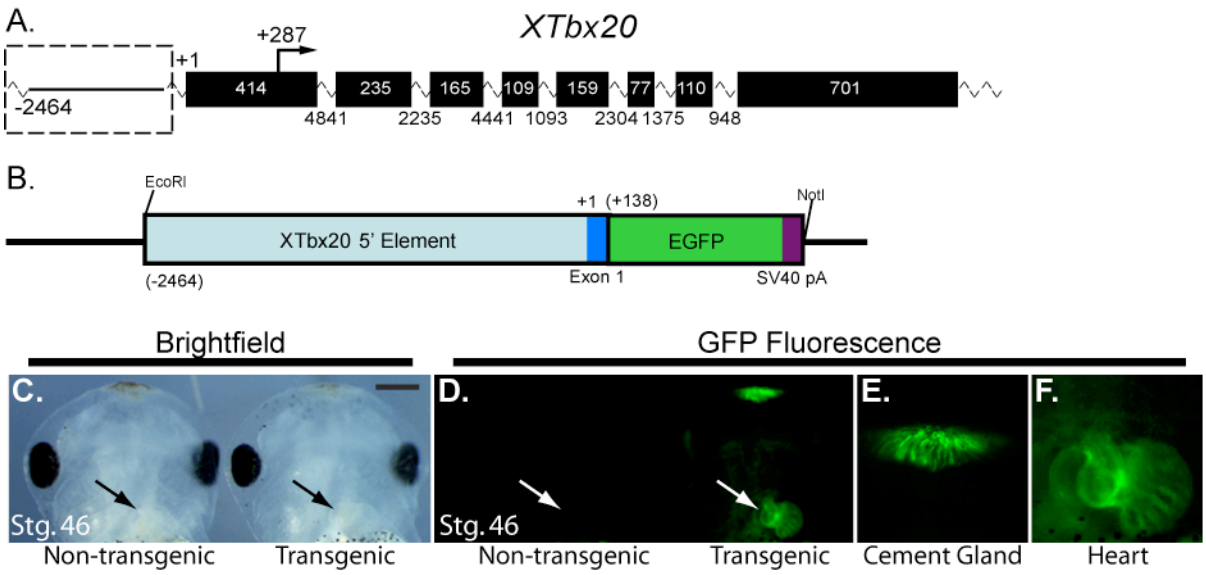


Figure 4.1. A regulatory element 5' to the *Tbx20* genomic locus is sufficient to drive gene expression in the cement gland and heart. A, Schematic representation of the *X. tropicalis Tbx20* genomic locus. *X. tropicalis Tbx20* consists of 8 exons spanning approximately 20kB. The *Tbx20* transcriptional start site is located 287bp upstream of the transcriptional start site in exon 1. A putative cardiac regulatory element is located at the 5' end of the *Tbx20* locus (dashed box). B, Schematic representation of the 2464bp region of the 5' end of *Tbx20* cloned in frame to the EGFP reporter to examine its regulatory capacity in *X. laevis* transgenics. C-F, EGFP expression is driven by 2464bp of the 5' end of *Tbx20* in the cement gland and heart of living *X. laevis* transgenic embryos. Arrows point to the heart. C, Ventral view of the anterior end of stage 46 sibling non-transgenic (left) or transgenic (right) embryos. D, Fluorescent view of siblings in C. E and F, Magnified view of the EGFP expression driven by the *Tbx20* regulatory element in the cement gland (E) and heart (F) of the transgenic embryo in D. Scale bar is equal to 500μm.

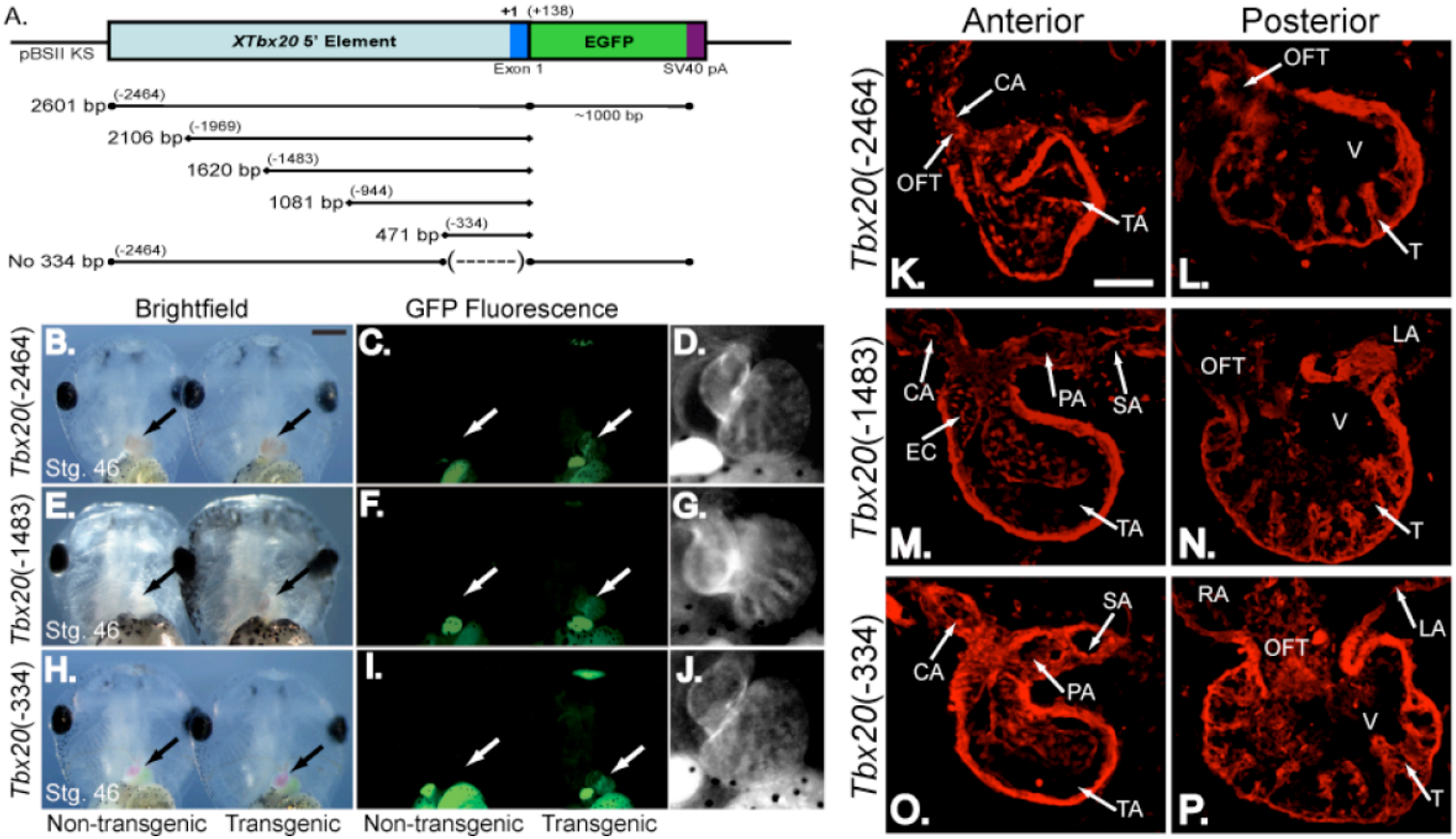

A.
+1 (+138)
XTbx20 5' Element
EGFP
pBSII KS
Exon 1
SV40 pA
2601 bp (-2464)
~1000 bp
2106 bp (-1969)
1620 bp (-1483)
1081 bp (-944)
471 bp (-334)
No 334 bp (-2464)
Brightfield
GFP Fluorescence
Tbx20(-2464)
Tbx20(-1483)
Tbx20(-334)
B.
C.
D.
E.
F.
G.
H.
I.
J.
Stg. 46
Non-transgenic
Transgenic
Non-transgenic
Transgenic
Anterior
Posterior
K.
L.
M.
N.
O.
P.
CA
OFT
TA
V
T
PA
SA
EC
LA
RA

Figure 4.2. A 334bp regulatory element recapitulates the endogenous expression of *Tbx20* throughout the *X. laevis* heart. A deletion series of the 5' regulatory element was created to determine a reduced element sufficient to drive EGFP transgene expression. A, Schematic representation of the deletion series of *Tbx20* elements fused to EGFP for *X. laevis* transgenesis. B, E, H, Ventral view of the anterior regions of living stage 46 (late tadpole) *X. laevis* embryos (left) and siblings transgenic for constructs shown in A (right) under white light. C, F, I Embryos as viewed under fluorescent light. Green autofluorescence in the gut can be noted in both control and transgenic embryos. Arrows note the location of the heart. Scale bar is equal to 500μm. D, G, I Magnified views of the EGFP-expressing hearts of embryos in C, F and I demonstrating that EGFP expression in the heart is maintained under the control of a *Tbx20*(-334) element. Transverse sections were cut through the embryos expressing *Tbx20*-EGFP shown in B-J, and expression of the *Tbx20(-2459)*-EGFP (K, L), *Tbx20(-1483)*-EGFP (M, N), and *Tbx20(-334)*-EGFP (O, P) transgenes is demonstrated by antibody staining for EGFP. Anterior (K, M, O) and posterior (L, N, P) sections show EGFP transgene expression throughout the heart. Scale bar is equal to 100μm. *TA – truncus arteriosis, OFT – outflow tract, CA – carotid arch, V – ventricle, T – trabeculae, EC – endocardial cushion, PA – pulmocutaneous arch, SA – systemic arch, LA – left atrium, RA – right atrium*

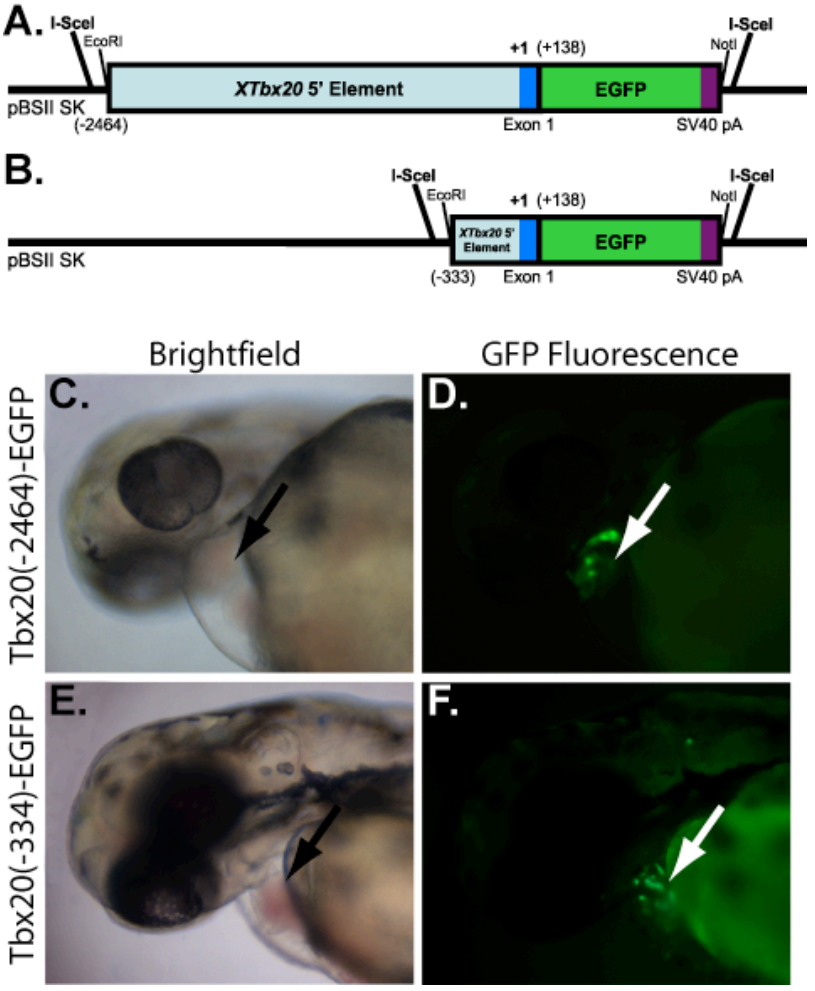


Figure 4.3. *X. laevis* regulatory elements of *Tbx20* are sufficient to drive EGFP reporter expression in the embryonic zebrafish heart. A, Schematic representation of the *XTbx20*-EGFP reporter constructs containing I-SceI meganuclease sites utilized in zebrafish. B, D Lateral views of live zebrafish embryos at 48 hpf. injected with the 2464bp *Tbx20*-EGFP transgene (B) or 334bp transgene (D). C, E Injected embryos express EGFP in the heart (arrows). Of embryos analyzed at 48 hpf, expression of EGFP in the heart was observed in 22% (59/264) of embryos injected with the 2464bp transgene, and in 20% (17/89) of embryos injected with the 334bp transgene. *hpf – hours post fertilization*

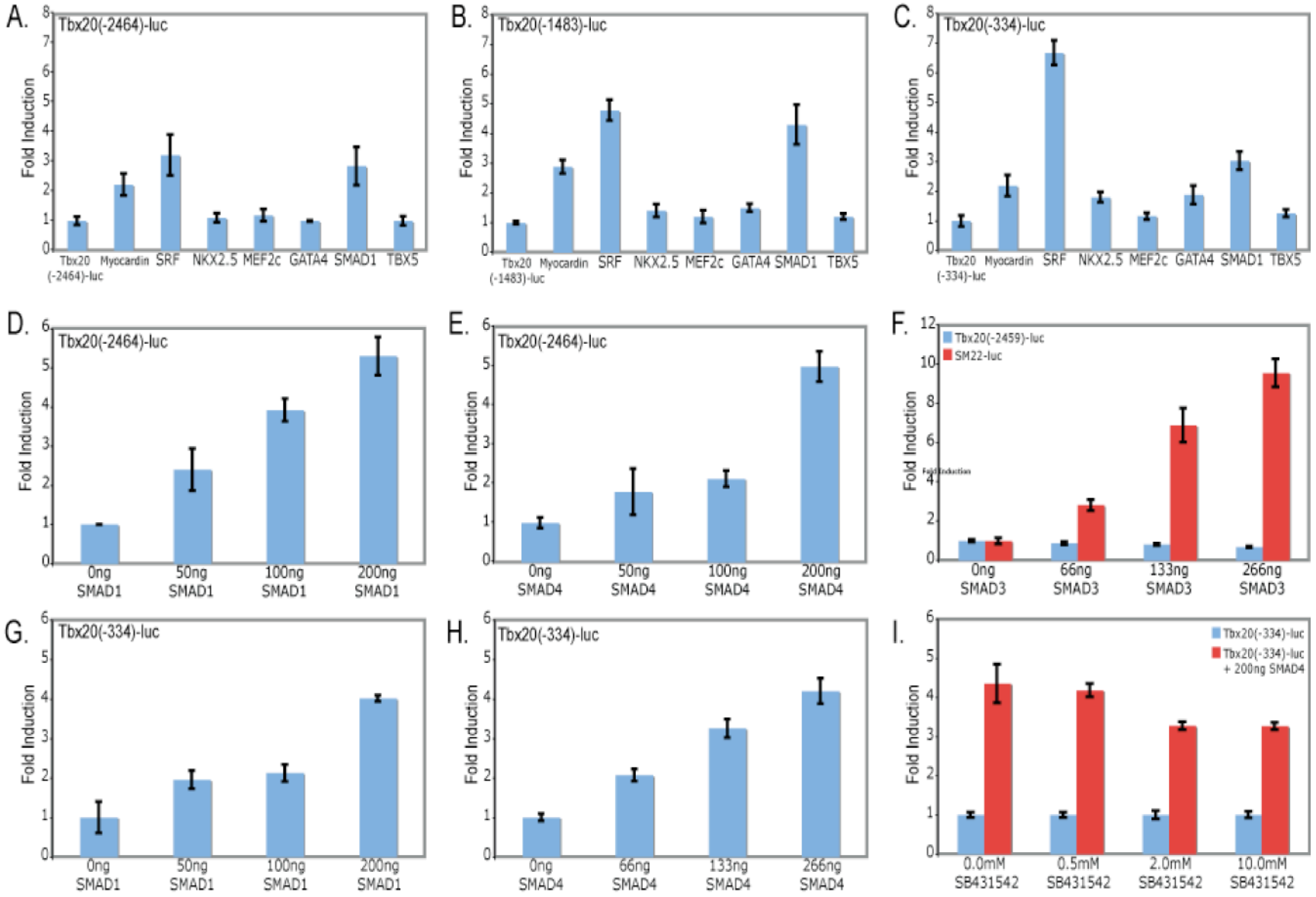

A.
Tbx20(-2464)-luc
Fold Induction
Tbx20 (-2464)-luc
Myocardin
SRF
NKX2.5
MEF2c
GATA4
SMAD1
TBX5
B.
Tbx20(-1483)-luc
Fold Induction
Tbx20 (-1483)-luc
Myocardin
SRF
NKX2.5
MEF2c
GATA4
SMAD1
TBX5
C.
Tbx20(-334)-luc
Fold Induction
Tbx20 (-334)-luc
Myocardin
SRF
NKX2.5
MEF2c
GATA4
SMAD1
TBX5
D.
Tbx20(-2464)-luc
Fold Induction
0ng SMAD1
50ng SMAD1
100ng SMAD1
200ng SMAD1
E.
Tbx20(-2464)-luc
Fold Induction
0ng SMAD4
50ng SMAD4
100ng SMAD4
200ng SMAD4
F.
Tbx20(-2459)-luc
SM22-luc
Fold Induction
0ng SMAD3
66ng SMAD3
133ng SMAD3
266ng SMAD3
G.
Tbx20(-334)-luc
Fold Induction
0ng SMAD1
50ng SMAD1
100ng SMAD1
200ng SMAD1
H.
Tbx20(-334)-luc
Fold Induction
0ng SMAD4
66ng SMAD4
133ng SMAD4
266ng SMAD4
I.
Tbx20(-334)-luc
Tbx20(-334)-luc + 200ng SMAD4
Fold Induction
0.0mM SB431542
0.5mM SB431542
2.0mM SB431542
10.0mM SB431542

Figure 4.4. *XTbx20* 5' regulatory elements are activated by TGF-β/BMP signaling via SMAD1 and SMAD 4 but not SMAD3. A-C, Luciferase reporters controlled by three *Tbx20* deletion elements were transfected into COS7 cells with a panel of cardiac factor expression plasmids. Both the *Tbx20*(-2464) (D-E) and *Tbx20*(-334) (G-H) reporters are activated by SMAD1 and SMAD4 in a dose-dependent manner when transfected with increasing amounts of SMAD expression plasmid. F, SMAD3 transfection does not induce the *Tbx20*(-2464) reporter, though the control SM22 reporter is dramatically induced. I, Treatment of COS7 cells with increasing doses of a small molecule inhibitor of activin signaling SB431542 does not affect the activation of the *Tbx20*(-334) plasmid by SMAD4. Values are the fold increase in luciferase activity relative to that driven by the reporter alone. Error bars represent the standard error of fold induction for three trials.

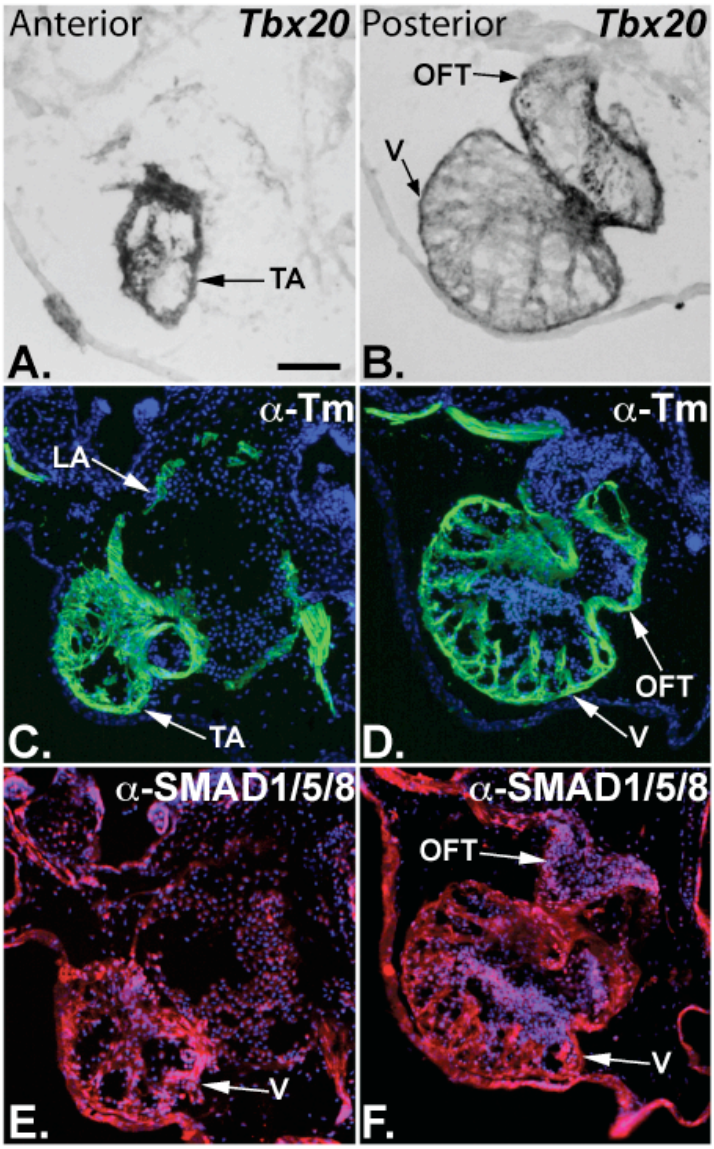
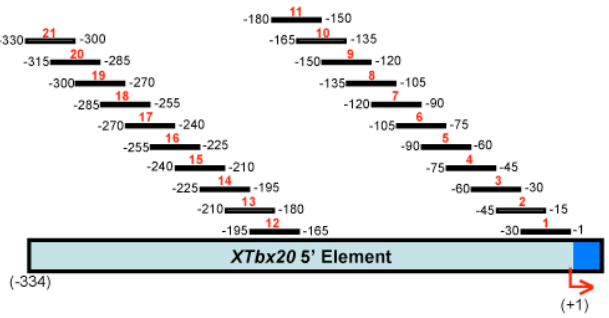


Figure 4.5. *XTbx20* is expressed throughout the myocardium and endocardium of the *X. laevis* heart. A and B, *Tbx20* is expressed in both the anterior and posterior regions of the *X. laevis* stage 46 heart. Serial sections show that *Tbx20* expression overlaps with that of the myocardial marker tropomyosin (C, D) and with SMAD1/5/8 expression in the endocardium (E, F) by immunohistochemistry. C-F, Anti-tropomyosin (Tm) staining is labeled in green, anti-SMAD1/5/8 is labeled in red, and all nuclei are labeled with DAPI in blue. Scale bar is equal to 100μm. *TA-truncus arteriosus, OFT-outflow tract, LA-left atrium, V-ventricle*

A.



C.

Oligo	Minimal SMAD Binding Element
Oligo 19	ATCCTTAGATTACTCT
Oligo 16	GTGCTTTTTTTTAGTA
Oligo 13	AAAGAAGTAGCTCGGC
Oligo 9	CATTCATTGCCTGTG
Oligo 8	CCCAGACACTTATCA
Oligo 6	CAGGCAGATGTGACGC
Oligo 2	CATGAAATTAAGGAAG

B.

OLIGO	Kd (μM)	Std. Dev.	Oligo Sequence
xVent	7.83	0.75	AGAGAGAATGTTTAGCATAACAATAGC
SRF Site	>7.42	0.64	AGCTTCTTTACACAGGATGTCCATATTAGGACATCTGCGTCAGCAA
Oligo 21	6.54	0.22	CTATTTGATCAGCAAACGAGATGGATTACA
Oligo 20	>7.42	0.36	ACGAGATGGATTACAGATGAGCATCCTTAG
Oligo 19*	3.76	0.22	GATGAGCATCCTTAGATTACTCTAAAAGCC
Oligo 18	>7.42	0.48	ATTACTCTAAAAGCCCCGCCCTTCTCTTAT
Oligo 17	5.63	0.41	CCGCCCTTCTCTTATGTCACGTGTGCTTTT
Oligo 16*	2.85	0.08	GTCACGTGTGCTTTTTTTTAGTAAGTCTTT
Oligo 15	>7.42	0.14	TTTTAGTAAGTCTTTTTCTCTCGGAGCCCA
Oligo 14	>7.42	0.77	TTCTCTCGGAGCCCAGTGAGAAAAAGAAGT
Oligo 13*	3.56	0.19	GTGAGAAAAAGAAGTAGCTCGGCTGATCCT
Oligo 12	>7.42	1.4	AGCTCGGCTGATCCTATCTGGCCCTGCTCC
Oligo 11	6.38	0.62	ATCTGGCCCTGCTCCATCCCTGCTGCCCTT
Oligo 10	5.97	0.60	ATCCCTGCTGCCCTTCATTCATTGCCTGTG
Oligo 9*	2.84	0.38	CATTCATTGCCTGTGCTCCAGCCGCCACCT
Oligo 8*	2.08	0.22	CTCCAGCCGCCACCTCCCAGACACTTATCA
Oligo 7	5.57	0.48	CCCAGACACTTATCAGCTGTATCAGGCAGA
Oligo 6*	2.40	0.18	GCTGTATCAGGCAGATGTGACGCTGCAGGG
Oligo 5	7.20	0.53	TGTGACGCTGCAGGGCTCCAATTGGCCAGG
Oligo 4	>7.42	0.33	CTCCAATTGGCCAGGAGAGAGATAGGATCT
Oligo 3	>7.42	0.49	AGAGAGATAGGATCTGTGTGGCCATGAAAT
Oligo 2*	3.11	0.10	GTGTGGCCATGAAATTAAGGAAGCAGAGGC
Oligo 1	>7.42	1.10	TAAGGAAGCAGAGGCTGAGAATGGGAACAG

D.

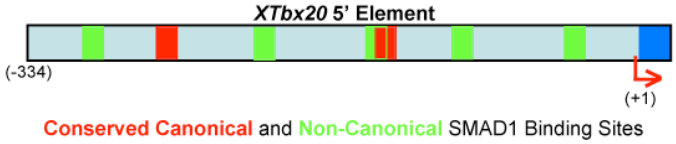


Figure 4.6. SMAD1 potentially binds to seven regions within the 334bp *Tbx20* regulatory element. A, Double stranded, 5' carboxyfluorescein-labeled, 30bp oligos designed for 2x coverage of the 334bp *Tbx20* cardiac regulatory element for use in fluorescence polarization assays. B, Dissociation constants (K_d), standard deviation of three trials and nucleotide sequence for each oligo analyzed in fluorescence polarization studies (*indicates oligos bound by SMAD1). K_d values for 9 oligos are recorded as >7.42μM, as K_d values significantly greater than the maximum SMAD1 concentration cannot be accurately calculated. C, Sequences of the minimal regions bound by SMAD1 on each of seven bound oligos. D, Schematic representation of the location of seven putative SMAD1 binding sites located within the 334bp cardiac regulatory element.

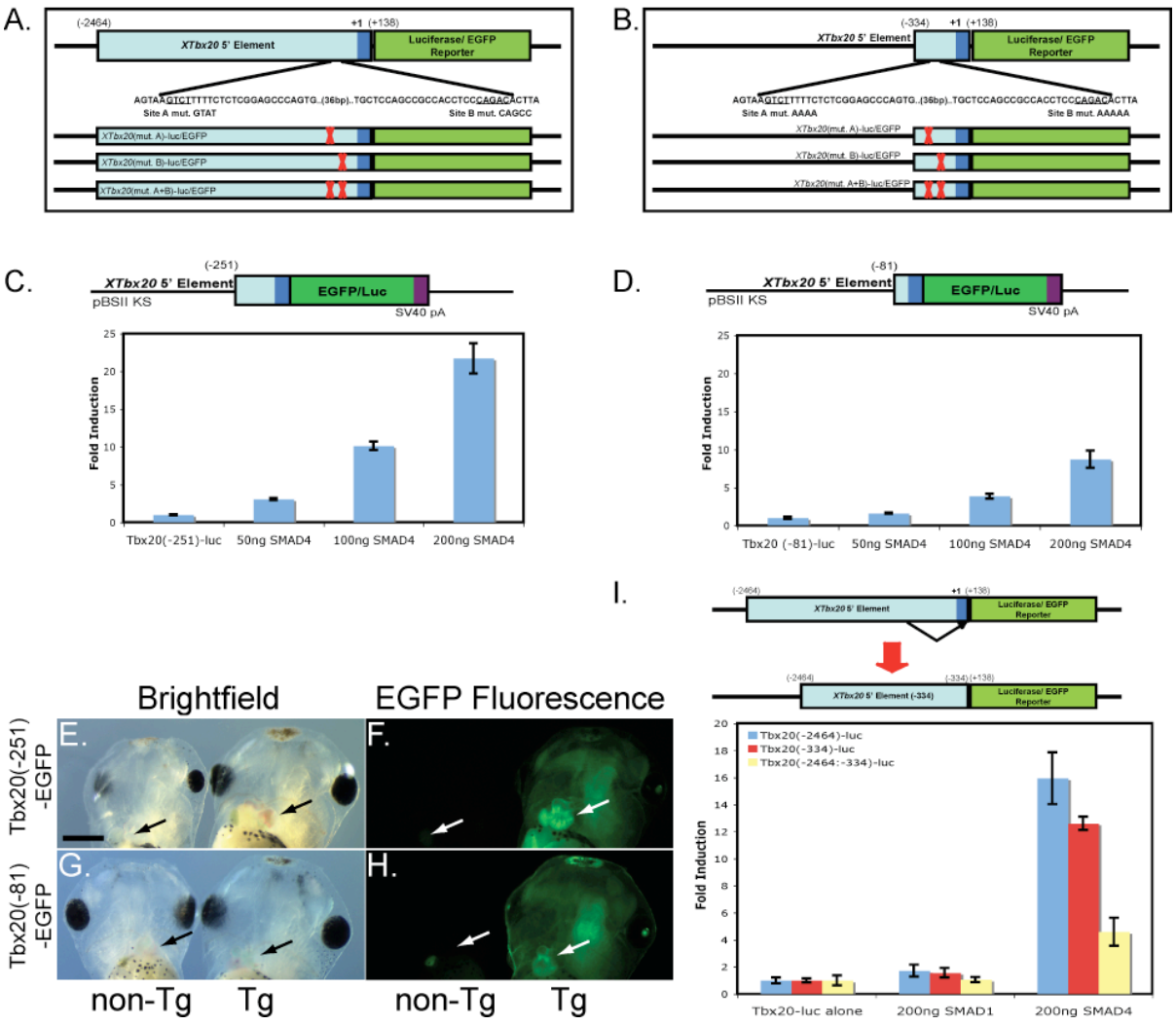

A.
(-2464)
+1 (+138)
XTbx20 5' Element
Luciferase/ EGFP Reporter
AGTAAGTCTTTTTCTCTCGGAGCCCAGTG..(36bp)..TGCTCCAGCCGCCACCTCCCAGACACTTA
Site A mut. GTAT
Site B mut. CAGCC
XTbx20(mut. A)-luc/EGFP
XTbx20(mut. B)-luc/EGFP
XTbx20(mut. A+B)-luc/EGFP
B.
(-334)
+1 (+138)
XTbx20 5' Element
Luciferase/ EGFP Reporter
AGTAAGTCTTTTTCTCTCGGAGCCCAGTG..(36bp)..TGCTCCAGCCGCCACCTCCCAGACACTTA
Site A mut. AAAA
Site B mut. AAAAA
XTbx20(mut. A)-luc/EGFP
XTbx20(mut. B)-luc/EGFP
XTbx20(mut. A+B)-luc/EGFP
C.
(-251)
XTbx20 5' Element
EGFP/Luc
pBSII KS
SV40 pA
Fold Induction
Tbx20(-251)-luc
50ng SMAD4
100ng SMAD4
200ng SMAD4
D.
(-81)
XTbx20 5' Element
EGFP/Luc
pBSII KS
SV40 pA
Fold Induction
Tbx20 (-81)-luc
50ng SMAD4
100ng SMAD4
200ng SMAD4
Brightfield
EGFP Fluorescence
E.
F.
G.
H.
Tbx20(-251) -EGFP
Tbx20(-81) -EGFP
non-Tg
Tg
non-Tg
Tg
I.
(-2464)
+1 (+138)
XTbx20 5' Element
Luciferase/ EGFP Reporter
(-2464)
(-334) (+138)
XTbx20 5' Element (-334)
Luciferase/ EGFP Reporter
Tbx20(-2464)-luc
Tbx20(-334)-luc
Tbx20(-2464:-334)-luc
Fold Induction
Tbx20-luc alone
200ng SMAD1
200ng SMAD4

Figure 4.7. Multiple SMAD1 binding sites are involved in, and necessary for *Tbx20* regulation. Two conserved SMAD binding sites were identified within the 334bp regulatory element using the ConSite algorithm and mutated using site-directed mutagenesis in the context of the *Tbx20*(-2464)-EGFP or –luc reporter constructs (A), and the *Tbx20*(-334)-EGFP or –luc constructs (B). C, COS7 cells were transfected with a truncated region of the *Tbx20* cardiac regulatory element retaining 251bp upstream of the transcriptional start site and increasing amounts of SMAD4 expression plasmid. D, COS7 cells were then transfected with the *Tbx20*(-81) plasmid containing only 81bp upstream of the transcriptional start and increasing amounts of SMAD4. Each additional deletion is activated by SMAD signaling in a dose-dependent manner. Values represent fold luciferase activity relative to reporter plasmid alone. E-H, *Tbx20*(-251)-EGFP (E-F) or *Tbx20*(-81)-EGFP (G-H) reporter constructs were introduced into *X. laevis* transgenic embryos. Transgenic embryos are located at the right of each image, while non-transgenic siblings are at the left. E and G, Brightfield views of living stage 46 embryos. F and H, EGFP expression of embryos in E and G. Arrows note the location of the heart. Scale bar is equal to 500μm. I, The 334bp regulatory element was eliminated from the 2464bp regulatory element to create *Tbx20*(-2464:-334)-luc and *Tbx20*(-2464:-334)-EGFP constructs. In COS7 cells, *Tbx20*(-2464)-luc, *Tbx20*(-334)-luc and *Tbx20*(-2464:-334)-luc were each co-transfected with 200ng Smad1 or Smad4 reporter constructs, demonstrating a reduction in *Tbx20* activation and a necessity for the 334bp upstream element. Fold induction reflects changes in induction relative to induction of the reporter alone, and error bars represent standard error of three replicates.

REFERENCES

1. **Ahn, D. G., I. Ruvinsky, A. C. Oates, L. M. Silver, and R. K. Ho.** 2000. *tbx20*, a new vertebrate T-box gene expressed in the cranial motor neurons and developing cardiovascular structures in zebrafish. Mech Dev **95:**253-258.

2. **Brown, C. O., 3rd, X. Chi, E. Garcia-Gras, M. Shirai, X. H. Feng, and R. J. Schwartz.** 2004. The cardiac determination factor, *Nkx2-5*, is activated by mutual cofactors GATA-4 and Smad1/4 via a novel upstream enhancer. J Biol Chem **279:**10659-10669.

3. **Brown, D. D., O. Binder, M. Pagratis, B. A. Parr, and F. L. Conlon.** 2003. Developmental expression of the *Xenopus laevis Tbx20* orthologue. Dev Genes Evol **212:**604-607.

4. **Brown, D. D., S. N. Martz, O. Binder, S. C. Goetz, B. M. J. Price, J. C. Smith, and F. L. Conlon.** 2005. Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development **132:**553-563.

5. **Cai, C. L., W. Zhou, L. Yang, L. Bu, Y. Qyang, X. Zhang, X. Li, M. G. Rosenfeld, J. Chen, and S. Evans.** 2005. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development **132**:2475-2487.

6. **Callis, T. E., D. Cao, and D. Z. Wang.** 2005. Bone morphogenetic protein signaling modulates myocardin transactivation of cardiac genes. Circ Res **97:**992-1000.

7. **Carson, C. T., E. R. Kinzler, and B. A. Parr.** 2000. *Tbx12*, a novel T-box gene, is expressed during early stages of heart and retinal development. Mech Dev **96:**137-140.

8. **Chang, P. S., L. Li, J. McAnally, and E. N. Olson.** 2001. Muscle specificity encoded by specific serum response factor-binding sites. J Biol Chem **276:**17206-17212.

9. **Clements, D., H. C. Taylor, B. G. Herrmann, and D. Stott.** 1996. Distinct regulatory control of the *Brachyury* gene in axial and non-axial mesoderm suggests separation of mesodermal lineages early in mouse gastrulation. Mech. Dev. **56:**139-149.

10. **Euler-Taimor, G., and J. Heger.** 2006. The complex pattern of SMAD signaling in the cardiovascular system. Cardiovasc Res **69:**15-25.

11. **Feng, X. H., X. Lin, and R. Derynck.** 2000. Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-beta. Embo J **19:**5178-5193.

12. **Gaussin, V., T. Van de Putte, Y. Mishina, M. C. Hanks, A. Zwijsen, D. Huylebroeck, R. R. Behringer, and M. D. Schneider.** 2002. Endocardial cushion and myocardial defects after cardiac myocyte-specific conditional deletion of the bone morphogenetic protein receptor ALK3. Proc Natl Acad Sci U S A **99:**2878-2883.

13. **Grabher, C., J. S. Joly, and J. Wittbrodt.** 2004. Highly efficient zebrafish transgenesis mediated by the meganuclease I-SceI. Methods Cell Biol **77:**381-401.

14. **Griffin, K. J., J. Stoller, M. Gibson, S. Chen, D. Yelon, D. Y. Stainier, and D. Kimelman.** 2000. A conserved role for H15-related T-box transcription factors in zebrafish and Drosophila heart formation. Dev Biol **218:**235-247.

15. **Heinritz, W., A. Moschik, A. Kujat, S. Spranger, H. Heilbronner, S. Demuth, A. Bier, M. Tihanyi, S. Mundlos, C. Gruenauer-Kloevekorn, and U. G. Froster.** 2005. Identification of new mutations in the TBX5 gene in patients with Holt-Oram syndrome. Heart **91:**383-384.

16. **Heldin, C., K. Miyazono, and P. ten Dijke.** 1997. TGF-β signaling from cell membrane to nucleus through SMAD proteins. Nature **390:**465-471.

17. **Hemmati-Brivanlou, A., and D. A. Melton.** 1992. A truncated activin receptor inhibits mesoderm induction and formation of axial structures in *Xenopus* embryos. Nature **359:**609-614.

18. **Henningfeld, K. A., S. Rastegar, G. Adler, and W. Knochel.** 2000. Smad1 and Smad4 are components of the bone morphogenetic protein-4 (BMP-4)-induced transcription complex of the *Xvent-2B* promoter. J Biol Chem **275:**21827-21835.

19. **Iio, A., M. Koide, K. Hidaka, and T. Morisaki.** 2001. Expression pattern of novel chick T-box gene, *Tbx20*. Dev Genes Evol **211:**559-562.

20. **Kaartinen, V., M. Dudas, A. Nagy, S. Sridurongrit, M. M. Lu, and J. A. Epstein.** 2004. Cardiac outflow tract defects in mice lacking ALK2 in neural crest cells. Development **131:**3481-3490.

21. **Kirk, E. P., M. Sunde, M. W. Costa, S. A. Rankin, O. Wolstein, M. L. Castro, T. L. Butler, C. Hyun, G. Guo, R. Otway, J. P. Mackay, L. B. Waddell, A. D. Cole, C. Hayward, A. Keogh, P. Macdonald, L. Griffiths, D. Fatkin, G. F. Sholler, A. M. Zorn, M. P. Feneley, D. S. Winlaw, and R. P. Harvey.** 2007. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. Am J Hum Genet **81:**280-291.

22. **Kraus, F., B. Haenig, and A. Kispert.** 2001. Cloning and expression analysis of the mouse T-box gene *tbx20*. Mech Dev **100:**87-91.

23. **Kretzschmar, M., and J. Massague.** 1998. SMADs: mediators and regulators of TGF-β signaling. Current Opinions in Genetics and Development **8:**103-111.

24. **Kroll, K. L., and E. Amaya.** 1996. Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signalling requirements during gastrulation. Development **122:**3173-3183.

25. **Kusanagi, K., H. Inoue, Y. Ishidou, H. K. Mishima, M. Kawabata, and K. Miyazono.** 2000. Characterization of a bone morphogenetic protein-responsive Smad-binding element. Mol Biol Cell **11:**555-565.

26. **Latinkic, B. V., M. Umbhauer, K. A. Neal, W. Lerchner, J. C. Smith, and V. Cunliffe.** 1997. The *Xenopus Brachyury* promoter is activated by FGF and low concentrations of activin and suppressed by high concentrations of activin and by paired-type homeodomain proteins. Genes Dev. **11:**3265-3276.

27. **Lerchner, W., B. V. Latinkic, J. E. Remacle, D. Huylebroeck, and J. C. Smith.** 2000. Region-specific activation of the *Xenopus brachyury* promoter involves active repression in ectoderm and endoderm: a study using transgenic frog embryos. Development **127:**2729-2739.

28. **Li, L., Z. Liu, B. Mercer, P. Overbeek, and E. N. Olson.** 1997. Evidence for serum response factor-mediated regulatory networks governing SM22alpha transcription in smooth, skeletal, and cardiac muscle cells. Dev Biol **187:**311-321.

29. **Liberatore, C. M., R. D. Searcy-Schrick, E. B. Vincent, and K. E. Yutzey.** 2002. *Nkx-2.5* gene induction in mice is mediated by a Smad consensus regulatory region. Dev Biol **244:**243-256.

30. **Lien, C. L., J. McAnally, J. A. Richardson, and E. N. Olson.** 2002. Cardiac-specific activity of an *Nkx2-5* enhancer requires an evolutionarily conserved Smad binding site. Dev Biol **244:**257-266.

31. **Liu, F., A. Hata, J. C. Baker, J. Doody, J. Carcamo, R. M. Harland, and J. Massague.** 1996. A human Mad protein acting as a BMP-regulated transcriptional activator [see comments]. Nature **381:**620-623.

32. **Maleki, S. J., C. A. Royer, and B. K. Hurlburt.** 2002. Analysis of the DNA-binding properties of MyoD, myogenin, and E12 by fluorescence anisotropy. Biochemistry **41:**10888-10894.

33. **Massague, J., J. Seoane, and D. Wotton.** 2005. Smad transcription factors. Genes Dev **19:**2783-2810.

34. **Mori, A. D., and B. G. Bruneau.** 2004. TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol **19:**211-215.

35. **Nieuwkoop, P. D., and J. Faber.** 1967. Normal Table of *Xenopus laevis* (Daudin). North Holland, Amsterdam.

36. **Oh, J., Z. Wang, D. Z. Wang, C. L. Lien, W. Xing, and E. N. Olson.** 2004. Target gene-specific modulation of myocardin activity by GATA transcription factors. Mol Cell Biol **24:**8519-8528.

37. **Prall, O. W., M. K. Menon, M. J. Solloway, Y. Watanabe, S. Zaffran, F. Bajolle, C. Biben, J. J. McBride, B. R. Robertson, H. Chaulet, F. A. Stennard, N. Wise, D. Schaft, O. Wolstein, M. B. Furtado, H. Shiratori, K. R. Chien, H. Hamada, B. L. Black, Y. Saga, E. J. Robertson, M. E. Buckingham, and R. P. Harvey.** 2007. An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation. Cell **128:**947-959.

38. **Qiu, P., X. H. Feng, and L. Li.** 2003. Interaction of Smad3 and SRF-associated complex mediates TGF-beta1 signals to regulate SM22 transcription during myofibroblast differentiation. J Mol Cell Cardiol **35:**1407-1420.

39. **Schlange, T., B. Andree, H. H. Arnold, and T. Brand.** 2000. BMP2 is required for early heart development during a distinct time period. Mech Dev **91:**259-270.

40. **Schneider, M. D., V. Gaussin, and K. M. Lyons.** 2003. Tempting fate: BMP signals for cardiac morphogenesis. Cytokine Growth Factor Rev **14:**1-4.

41. **Schultheiss, T. M., J. B. Burch, and A. B. Lassar.** 1997. A role for bone morphogenetic proteins in the induction of cardiac myogenesis. Genes Dev **11:**451-462.

42. **Schwartz, R. J., and E. N. Olson.** 1999. Building the heart piece by piece: modularity of *cis*-elements regulating *Nkx2-5* transcription. Development **126:**4187-4192.

43. **Showell, C., K. S. Christine, E. M. Mandel, and F. L. Conlon.** 2006. Developmental expression patterns of *Tbx1*, *Tbx2*, *Tbx5*, and *Tbx20* in *Xenopus tropicalis*. Dev Dyn **235:**1623-1630.

44. **Singh, M. K., V. M. Christoffels, J. M. Dias, M. O. Trowe, M. Petry, K. Schuster-Gossler, A. Burger, J. Ericson, and A. Kispert.** 2005. Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. Development **132:**2697-2707.

45. **Smith, J. C., B. M. Price, J. B. A. Green, D. Weigel, and B. G. Herrmann.** 1991. Expression of a *Xenopus* homolog of *Brachyury* (T) is an immediate-early response to mesoderm induction. Cell **67:**79-87.

46. **Staehling Hampton, K., F. M. Hoffmann, M. K. Baylies, E. Rushton, and M. Bate.** 1994. dpp induces mesodermal gene expression in *Drosophila*. Nature **372:**783-786.

47. **Stennard, F. A., M. W. Costa, D. Lai, C. Biben, M. B. Furtado, M. J. Solloway, D. J. McCulley, C. Leimena, J. I. Preis, S. L. Dunwoodie, D. E. Elliott, O. W. Prall, B. L. Black, D. Fatkin, and R. P. Harvey.** 2005. Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. Development **132**:2451-2462.

48. **Stott, D., A. Kispert, and B. G. Herrmann.** 1993. Rescue of the tail defect of Brachyury mice. Genes Dev **7:**197-203.

49. **Szeto, D. P., K. J. Griffin, and D. Kimelman.** 2002. HrT is required for cardiovascular development in zebrafish. Development **129:**5093-5101.

50. **Takeuchi, J. K., M. Mileikovskaia, K. Koshiba-Takeuchi, A. B. Heidt, A. D. Mori, E. P. Arruda, M. Gertsenstein, R. Georges, L. Davidson, R. Mo, C. C. Hui, R. M. Henkelman, M. Nemer, B. L. Black, A. Nagy, and B. G. Bruneau.** 2005. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. Development **132**:2463-2474.

51. **Wang, D., P. S. Chang, Z. Wang, L. Sutherland, J. A. Richardson, E. Small, P. A. Krieg, and E. N. Olson.** 2001. Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. Cell **105:**851-862.

52. **Wang, D. Z., S. Li, D. Hockemeyer, L. Sutherland, Z. Wang, G. Schratt, J. A. Richardson, A. Nordheim, and E. N. Olson.** 2002. Potentiation of serum response factor activity by a family of myocardin-related transcription factors. Proc Natl Acad Sci U S A **99:**14855-14860.

53. **Wang, D. Z., M. R. Valdez, J. McAnally, J. Richardson, and E. N. Olson.** 2001. The *Mef2c* gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. Development **128:**4623-4633.

54. **Weinberg, R. L., D. B. Veprintsev, and A. R. Fersht.** 2004. Cooperative binding of tetrameric p53 to DNA. J Mol Biol **341:**1145-1159.

55. **Xu, X., Z. Yin, J. B. Hudson, E. L. Ferguson, and M. Frasch.** 1998. Smad proteins act in combination with synergistic and antagonistic regulators to target Dpp responses to the *Drosophila* mesoderm. Genes Dev **12:**2354-2370.

56. **Yang, L., C. L. Cai, L. Lin, Y. Qyang, C. Chung, R. M. Monteiro, C. L. Mummery, G. I. Fishman, A. Cogen, and S. Evans.** 2006. Isl1-Cre reveals a common Bmp pathway in heart and limb development. Development **133:**1575-1585.

CHAPTER 5

THE ROLE OF microRNA-1 AND -133 IN SKELETAL MUSCLE PROLIFERATION AND DIFFERENTIATION

PREFACE

Chapter 5 focuses on the regulation and roles of two members of a newly characterized class of regulatory RNAs termed microRNAs, members of which are directly involved in the post-transcriptional regulation of gene expression. Here, miRNA-1 and miRNA-133 are shown to be transcribed from the same loci while performing unique functions in early skeletal muscle proliferation and differentiation, the first evidence of its kind. This suggests that miRNAs are key members of a transcriptional network that is involved in early muscle and embryonic development.

This work was previously published and represents a collaboration with Jian-Fu Chen, J. Michael Thomson, Qiulian Wu, Thomas E. Callis, Scott M. Hammond and Da-Zhi Wang. The project was designed and developed in the lab of Dr. Da-Zhi Wang by Jian-Fu Chen and Dr. Da-Zhi Wang. To contribute to this endeavor, I carried out the majority of the *Xenopus* studies and was an active participant in the writing of the manuscript and critical.

Chen, J.F., E.M. Mandel, J.M. Thomson, Q. Wu, T.E. Callis, S.M. Hammond, F. L. Conlon, and D.Z. Wang. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nature Genetics **38(2)**:228-233.

SUMMARY

Understanding the molecular mechanisms that regulate cellular proliferation and differentiation is a central theme of developmental biology. MicroRNAs (miRNAs) are a recently discovered class of ~22-nucleotide regulatory RNAs that post-transcriptionally regulate gene expression (1, 2). Increasing evidence has pointed to the potential role of miRNAs in a variety of biological processes (6, 8, 9, 15, 27, 28). Here, we describe that miRNA-1 (miR-1) and miRNA-133 (miR-133), which are clustered on the same chromosomal loci, are transcribed together in a tissue-specific manner during development. miR-1 and miR-133 play distinct roles in modulating skeletal muscle proliferation and differentiation in cultured myoblasts *in vitro* and in *Xenopus* embryos *in vivo*. miR-1 promotes myogenesis by targeting histone deacetylase 4 (HDAC4), a transcriptional repressor of muscle gene expression. In contrast, miR-133 enhances myoblast proliferation by repressing serum response factor (SRF). The results reveal, for the first time, that two mature miRNAs, derived from the same miRNA polycistron and transcribed together, could perform distinct biological functions. Together, our studies suggest a molecular mechanism in which miRNAs participate in transcriptional circuits that control skeletal muscle gene expression and embryonic development.

BACKGROUND, RESULTS AND DISCUSSION

In order to understand the potential involvement of microRNAs (miRNAs) in skeletal muscle proliferation and differentiation, we analyzed the expression of miRNAs during skeletal muscle differentiation using the established microarray analysis (23). We chose to use C2C12 myoblasts because this line of cells faithfully mimics skeletal muscle

differentiation *in vitro* as myoblasts can be induced to become terminally differentiated myotubes when serum is withdrawn from the culture medium (3, 18, 22). We found that the expression of a fraction of the miRNAs examined was up-regulated in differentiated C2C12 myoblasts/myotubes (Fig. 5.1a and Fig. S.5.1). The increase in expression of miR-1 and miR-133 in differentiated myoblasts was confirmed by Northern blot analysis (Fig. 5.1b and Fig. S.5.2).

miR-1 and miR-133 are specifically expressed in adult cardiac and skeletal muscle tissues, but not in other tissues tested (Fig. 5.1c, Fig. 5.3)(13-15). However, little is known about the temporospatial distribution of specific miRNAs during mammalian development. We therefore examined the expression of miR-1 and miR-133 in mouse embryos and neonates. miR-1 and miR-133 are expressed at very low levels in the developing hearts and skeletal muscle of E13.5 and E16.5 embryos, (Fig. 5.1d and Fig. S.5.3). An increasing level of miR-1 and miR-133 expression was found in neonatal hearts and skeletal muscle, though it is still significantly lower than that of adults (Fig. 5.1e and Fig. S.5.3). These data are consistent with findings from zebrafish in which the majority of miRNAs are expressed relatively late during embryogenesis (26).

Both miR-1 and miR-133 are clustered together on mouse chromosomes 2 (separated by 9.3 kb) and 18 (separated by 2.5 kb) (Fig. S.5.4)(14). We performed a Northern blot analysis using ~300 bp genomic probes including the miR-1 or miR-133 sequences (Fig. S.5.4). miR-1 and miR-133 probes from chromosome 18 detected a single primary transcript of ~ 6 kb from total RNAs isolated from heart and skeletal muscle (Fig. S.5.4 b and c), indicating that miR-1 and miR-133 are indeed transcribed together. While both miR-1 and miR-133 probes from chromosome 2 detected a transcript of ~ 10 kb from the heart and

skeletal muscle, the miR-133 probe also hybridized to two additional transcripts of ~ 4.5 kb and ~ 2.2 kb, while the miR-1 probe also detected a major transcript of ~ 6 kb (Fig. S.5.4 d and e), suggesting the potential involvement of post-transcriptional processing. Together, our data indicate that cardiac- and skeletal muscle-specific expression of miR-1 and miR-133 is dictated at the primary transcription step.

We reasoned that the regulatory elements which control the transcription of both chromosome 2, and 18 miR-1 and miR-133 clusters are likely conserved. We therefore performed sequence analysis and identified a highly conserved region (~ 2 kb) which lies about 50 kb upstream of the miR-1/133 clusters on both chromosome 2 and 18 (Fig. S.5.5). When this genomic fragment from chromosome 2 was used to drive the expression of a dsRed reporter gene in transgenic *Xenopus*, we found cardiac- and skeletal-muscle specific expression of the transgene (Fig. S.5.5).

To assess the function of miR-1 and miR-133 in skeletal muscle, we first attempted to overexpress miR-1 and miR-133 in mammalian cells. We tested and validated the expression and activity of both miRNAs using Northern blot analysis as well as miR-1 and miR-133 "sensors"(19), in which the complementary sequences for miR-1 or miR-133 were cloned downstream of a dsRed coding sequence (Fig. S.5.6 and Data not shown). We transfected C2C12 myoblasts with miR-1 or miR-133 and then either maintained cells in growth medium (GM) or transferred them to differentiation medium (DM) after transfection. miR-1 strongly enhanced myogenesis as indicated by increased expression of both the early and late myogenic markers myogenin and myosin heavy chain (MHC), respectively, as well as other myogenic markers, including MyoD, Mef2, and skeletal α-actin (Fig. 5.2, a-e, i, j and Table 5.1). miR-1 induced myogenic marker gene expression in cells maintained in both the log-

phase growth condition (Fig. 5.2c) and the differentiation condition (Fig. 5.2, d, e). Accelerated myogenic differentiation induced by miR-1 is also accompanied by a decrease in cell proliferation, as marked by a significant decrease in the expression of phospho-histone H3 (Fig. 5.2, c, e and Table 5.1). Of particular note, miR-1 induced myogenesis is specific, since overexpression of a GFP control or miR-208, which is not endogenously expressed in skeletal muscle cells, showed no effect (Fig. 5.2 a-e). Furthermore, mutations introduced into miR-1 "seed" sequences abolished its ability to activate myogenic gene expression (Fig. 5.2, d, e). In contrast, overexpression of miR-133 repressed the expression of myogenin and MHC (Fig. 5.2, a-e and Table 5.1) and promoted myoblast proliferation (Fig. 5.2, c, e and Table 5.1). Again, the effect of miR-133 on myoblasts proliferation is specific, as controls showed no effect and mutation introduced abolished the function of miR-133 (Fig. 5.2, a-e, j).

We performed the reciprocal experiment wherein we transfected C2C12 myoblasts with the miR-1 or miR-133 2'-O-methyl antisense inhibitory oligos (or control GFP and miR-208), which have been shown to inhibit the function of miRNAs(10, 21). Cells transfected with the miR-1 inhibitor showed inhibition of myogenesis and promotion of myoblast proliferation, as indicated by a decrease in myogenic markers and an increase in phospho-histone H3 (Fig. 5.2, f-i and Table 5.1). Consistent with the role of miR-133 in promoting myoblast proliferation and repressing differentiation, inhibition of miR-133 caused an opposing effect, where myogenesis was enhanced and cell proliferation repressed (Fig. 5.2, f-j and Table 5.1). In contrast, control 2'-O-methyl inhibitors showed no effects (Fig. 5.2, f-j). We conclude that miR-1 and miR-133 have distinct roles in skeletal muscle

proliferation and differentiation: miR-1 promotes myoblast differentiation, whereas miR-133 stimulates myoblast proliferation.

Both miR-1 and miR-133 have been found in most animal species, from *Drosophila* to human, suggesting they are evolutionary conserved. To test the effects of miR-1 and miR-133 on skeletal muscle and heart development *in vivo*, we identified copies of miR-1 and miR-133 in *Xenopus* and tested their function through mis-expression. Introduction of miR-1 at the one cell stage leads to a dramatically shortened axis with accompanying reduction in anterior structures and an increase in body size along the dorsal-ventral axis compared to either uninjected or miGFP injected controls ($n > 45$, two independent experiments) (Fig. 5.3). Although somites formed in miR-1 injected embryos (Fig. 5.3), whole-mount antibody staining and serial sectioning reveal the tissue is highly disorganized and fails to develop into segmented structures (Fig. 5.3, e, f, j). Cardiac tissue is completely absent as judged by histology, tropomyosin staining (Fig. 5.3, f, j) and cardiac actin staining (Data not shown). In addition to these defects, there is a dramatic decrease in phospho-histone H3 staining (Fig. 5.3, i-k), consistent with the essential role of miR-1 in regulating muscle cell proliferation and differentiation. Although mis-expression of miR-133 also leads to a reduction in anterior structures and defects in somite development, in contrast to miR-1, there is only a modest reduction in anterior-posterior length and somatic defects are most severe in the more anterior or posterior aspects of the embryo where somites fail to form (Fig. 5.3, g, h). In addition, cardiac tissue frequently forms in miR-133 embryos, though it is highly disorganized and fails to undergo cardiac looping or chamber formation (Fig. 5.3, g, h, k and Data not shown). Collectively, these data suggest that the correct timing and levels of both miR-1 and miR-133 are required for proper skeletal muscle and heart development.

To identify target genes that could mediate the observed effects of miR-1 and miR-133 on skeletal muscle proliferation and differentiation, we examined predicted target genes for these two miRNAs. Multiple computational/bioinformatics-based approaches have been applied to predict targets for miRNAs and numerous potential targets have been suggested (11, 12, 16). Strikingly, many transcription factors have been suggested to be the targets for miRNAs, raising the possibility that miRNAs might be involved in transcriptional regulation of gene expression. Among the predicted miR-1 targets, HDAC4 has previously been shown to inhibit muscle differentiation and skeletal muscle gene expression, mainly by repressing MEF2C, an essential muscle-related transcription factor (18, 20). HDAC4 contains two naturally occurring putative miR-1 sites at its 3' UTR, which are evolutionarily conserved among vertebrate species (Fig. S.5.7). Similarly, two conserved miR-133 binding sites are found in the 3' UTR of the mammalian SRF gene (Fig. S.5.7), which has been shown to play an important role in muscle proliferation and differentiation *in vitro* and *in vivo* (17, 22, 25).

We fused the 3' UTRs of mouse SRF and HDAC4 to a luciferase reporter gene and transfected these constructs along with transfection controls into mammalian cells. Ectopic overexpression of miR-1 strongly repressed a HDAC4 3' UTR luciferase reporter gene, whereas miR-133 inhibited the expression of a SRF 3' UTR luciferase reporter gene (Fig. 5.4a). In contrast, mutations introduced into miR-1 or miR-133 "seed" sequences abolished such repression, indicating the specificity of the action (Fig. 5.4a).

When the above reporters were transfected into C2C12 myoblasts and luciferase activity measured before and after the induction of cell differentiation, we found that the reporter activity was dramatically repressed in differentiated cells (Fig. 5.4b), indicating that increased levels of endogenous miR-1 and miR-133 inhibited the reporter gene. The effects

and specificity of endogenous miR-1 and miR-133 were monitored by the miRNA "sensor" (Fig. S.5.6). In contrast, the luciferase activity of the MCK-luc reporter, an indicator of muscle differentiation, was increased in differentiated muscle cells (Fig. 5.4b). Furthermore, overexpression of miR-1 led to the down-regulation of endogenous HDAC4 protein in C2C12 cells in both the growth condition (Fig. 5.4c) and differentiation condition (Fig. 5.4e), whereas miR-133 repressed the expression of endogenous SRF proteins (Fig. 5.4, c, e). In contrast, the mRNA levels of SRF and HDAC4 were not altered by those miRNAs (Fig. 5.4d), supporting the notion that miRNAs repress the function of their target genes mainly by inhibiting translation. When 2'-O-methyl-antisense oligos against miR-1 or miR-133 were applied, they relieved repression exerted on the protein levels of HDAC4 or SRF, respectively (Fig. 5.4g), with no effect on their mRNA levels (Fig. 5.4f). To further verify that HDAC4 and SRF are cognate targets for miR-1 or miR-133 in regulating skeletal muscle gene expression, we tested whether co-transfecting expression plasmids for SRF or HDAC4 could "suppress" miRNA-mediated myogenesis. Indeed, as shown in Fig. 5.4h, overexpression of SRF partially reversed myogenic gene repression induced by miR-133. In contrast, HDAC4 counteracted the effects of miR-1 on skeletal muscle gene expression (Fig. 5.4h). Consistent with the potential involvement of HDAC4 and SRF in miR-1 and mir-133-dependent skeletal muscle proliferation and differentiation, endogenous HDAC4 and SRF protein levels were down-regulated in differentiated C2C12 cells, with a concomitant increase in expression of myogenic differentiation markers and a decrease in expression of the mitotic index marker phospho-histone H3 (Fig. 5.4i and Fig. S.5.2d). Decreased expression of SRF and HDAC4 proteins was accompanied by an increase expression of miR-1 and miR-133 (compare Fig. 5.4i with Fig. 5.1b). Together, these data demonstrate that

miR-1 and miR-133 specifically repress HDAC4 and SRF protein levels, respectively, which in turn, contributes to (at least in part) the regulatory effects of those miRNAs on myoblast proliferation and differentiation.

We characterized cardiac- and skeletal muscle-specific miR-1 and miR-133 and have shown their essential function in controlling skeletal muscle proliferation and differentiation. Most importantly, we found that miR-1 and miR-133, which are clustered on the same chromosomal loci and transcribed together as a single transcript, become two independent, mature miRNAs with distinct biological functions achieved by inhibiting different target genes. This implicates the involvement of miRNAs in complex molecular mechanisms. Interestingly, while the tissue-specific expression of miR-1 and miR-133 is controlled by myoD and SRF (28), SRF expression is repressed by miR-133. Therefore, these findings reveal a negative regulatory loop in which miRNAs participate in regulatory pathways to control cellular proliferation and differentiation (Fig. 5.5). It will be interesting, for the future, to determine whether miR-1 and miR-133 are involved in cardiac- and skeletal muscle-related human diseases.

MATERIALS AND METHODS

MicroRNA expression analysis by microarray

Total RNA was isolated from C2C12 cells cultured in growth medium (GM) consisting of DMEM (Sigma) with 10% fetal bovine serum (FBS) (Sigma) and 1% penicillin/streptomycin (Invitrogen) or differentiation medium (DM) consisting of DMEM (Sigma) with 2% horse serum (Sigma) at different time points (day 0, 1, 3, and 5 with the first day transferring into DM counted as day 0). Microarray hybridization was performed and data analyzed as

described(23). Briefly, 2.5 ug isolated RNA was labeled with 5'-phosphate-cytidyl-uridyl-Cy3-3' (Dharmacon) using RNA ligase and hybridized with 0.5 mM mixture of oligonucleotide probes for 124 microRNAs labeled with Alexa 647 (Cy5) in disposable chambers from MJ Research (part number SLF-0601). Normalized log (base 2) data was hierarchically clustered by gene and is plotted as a heat map. The range of signal was from –4 fold to +4 fold. Yellow denotes high expression and blue denotes low expression, relative to the median.

Northern blot analysis

Total RNA was extracted from C2C12 cells, mouse embryonic or adult tissue using Trizol Reagent (Invitrogen). For Northern blot analysis of miRNA, PEG was applied to remove large sized RNAs. Briefly, 30 μg of each total RNA sample were mixed 1:1 with 5X PEG solution and placed on ice 10 min. After 10 min centrifuging at maximum speed at 4°C, the supernatant was transferred to a fresh tube. RNAs were then precipitated by adding 2.5 volumes of 100% EtOH and centrifuged 30 min at maximum speed. Northern blot analysis for miRNAs was performed as described (14). miR-1 and miR-133 oligonucleotide sequences used as probes were listed in Table S.5.1. Northern blot analysis was used to detect primary transcripts of miRNAs and performed as described(24), using 20 μg of total RNA from each sample. Genomic fragments for miR-1 and miR-133 were PCR-cloned and serve as probes.

Cloning and expression of miR-1 and miR-133

Genomic fragments for miR-1 and miR-133 precursors from mouse chromosomes 2 and 18 (ch 2 or ch 18) were PCR amplified using mouse genomic DNA as a template (for PCR primers, see Table S.5.1). The PCR products were cloned into the pDNA(+)3.1 vector (Invitrogen) and the expression of miRNAs was determined by transfecting expression vectors into mammalian cells (COS7, 10T1/2 or C2C12) and detected by Northern blot analysis.

Cell culture, *in vitro* myogenesis differentiation and luciferase reporter assay

C2C12 myoblast cells were cultured and myogenesis induced as described (18). Transient transfection luciferase reporter assays were performed as described (18, 24).miRNA duplexes and 2'-*O*-methyl antisense oligoribonucleotides for miR-1, miR-133, miR-208 and GFP were purchased from Dharmacon (see Table S.5.1 for sequences). They were introduced into mammalian cells using either Lipofectamin (Invitrogen) transfection (200 nM) or eletroporation using Amaxa Nucleofector system (5 μg).

For 3' UTR-luciferase reporter construction, the multiple cloning site of the pGL3-Control Vector (Promega) was removed and placed downstream of the luciferase gene. 3' UTRs for mouse *HDAC4* and *SRF* were PCR amplified and cloned into the modified pGL3-Control Vector to result in the constructs SRF-3'UTR and HDAC4-3'UTR (see Table S.5.1 for PCR primer sequences). Luciferase reporter assays were performed as described (24).

Western blot and immunostaining

Western blots were performed as described (5). The following antibodies were used: Anti-myogenin, SRF, MEF2, HDAC4 and β-tubulin (Santa Cruz Biotechnology, Santa Cruz, CA); phospho-histone H3 (Upstate Biotechnology, Lake Placid, NY). The MF20 antibody, which recognizes striated muscle-specific MHC, was obtained from the DSHB (University of Iowa, Iowa City, IA). For immunostaining, treated C2C12 cell in 12-well plates were fixed with 4% formaldehyde for 5 min at 37°C and changed to 0.1% NP40/PBS solution for 15 min at RT. Primary antibodies were incubated in 0.1% NP40-PBS with 3% BSA for 2 hr in the following concentration: anti-myogenin (1: 20 dilution), anti-phospho-histone H3 (1:100 dilution), MF20 (1:10 dilution). Secondary antibodies fluorescein anti-mouse/rabbit (1:100 dilution, Vector laboratories, Inc) were adding in 0.1% NP40-PBS with 3% BSA for 1 hr at 37°C. DAPI was added in for 5 min at RT. After several wash with PBS, cells were subjected to fluorescence microscopy observation. Ten fields that cover the whole well were picked and green fluorescence positive cells and total cells with DAPI staining were counted for each well, respectively.

RT-PCR analysis

RT-PCR was performed essentially as described (5). Total RNA were extracted from C2C12 cells using Trizol Reagent (Invitrogen), and 2.0 μg aliquots were reverse transcribed to cDNA using random hexamers and MMLV reverse transcriptase (Invitrogen). For each case, 2.5% cDNA pool was used for amplification and PCR were performed for 24–28 cycles. Sequences for PCR primers can be found in Table S.5.1.

Xenopus embryo injections and transgenesis

Standard methods were utilized in obtaining and culturing *Xenopus laevis* embryos. DNA constructs were linearized with Kpn I and transgenic embryos were generated according to the methods described by Kroll & Amaya (13). Expression of the transgene was analyzed under a Leica MZFLIII microscope. Preparation and injection of *Xenopus* with miRNAs was carried out essentially as previously described (7). However, RNA was not capped prior to injection. Whole-mount immunohistology analysis was carried out as described (4).

ACKNOWLEDGEMENTS

We thank members of our laboratories for discussion and support. We would also like to thank Tim McKinsey for advice on the HDAC4 antibodies, Enrique Amaya for the dsRed construct, Susan Smyth for allowing us to use her electroporation apparatus, and Morgan von Drehl for histology. We thank Drs. Mark Majesky and Cam Patterson for stimulating discussion and critical reading of the manuscript. E.M.M. is funded by a NSF graduate research fellowship. J.M.T. is a Frederick Gardner Cottrell Postdoctoral Fellow. S.M.H. is a General Motors Cancer Research Foundation Scholar. F.L.C. was supported by NIH and the American Heart Association. D.-Z.W. is a Basil O'Connor Scholar of the March of Dimes Birth Defects Foundation and was supported by NIH, the Muscular Dystrophy Association and an American Heart Association Grant-in-Aid.

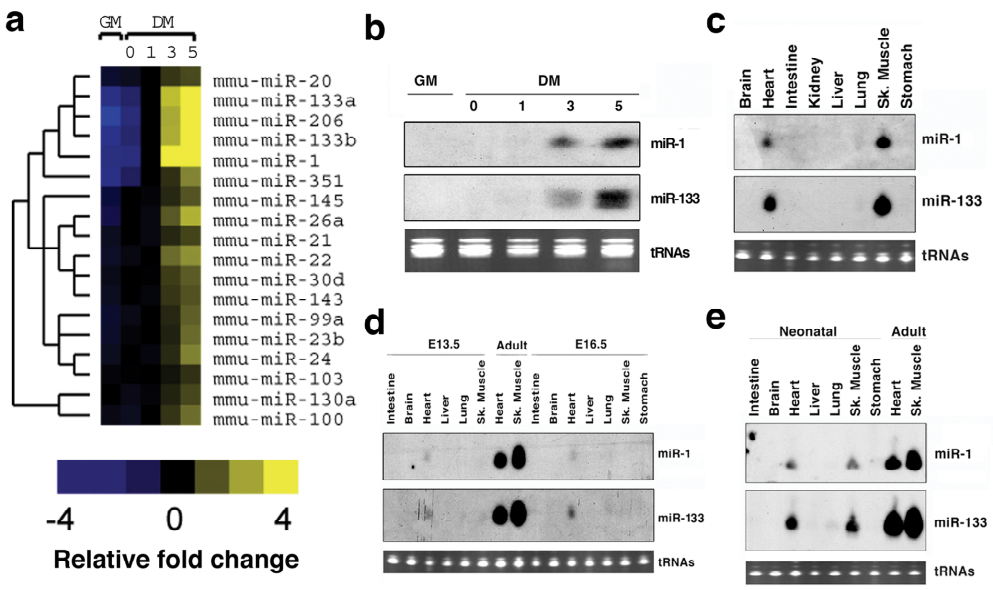


Figure 5.1. Expression of miR-1 and miR-133 in cardiac and skeletal muscle during development. (**a**) miRNA array expression data from C2C12 myoblasts cultured in growth medium (GM) or in differentiation medium (DM) for 0, 1, 3 and 5 days, respectively. Normalized log (base 2) data was hierarchically clustered by gene and is plotted as a heat map. The range of signal was from –4 fold to +4 fold. Yellow denotes high expression and blue denotes low expression, relative to the median and only the miRNA nodes that are up-regulated in differentiation medium are shown. (**b**) Northern blot analysis of the expression of miR-1 and miR-133 using total RNA isolated from C2C12 myoblasts cultured in GM or in DM for 0, 1, 3 and 5 days, respectively. tRNAs were used as a loading control. (**c**) Northern blot analysis of the expression of miR-1 and miR-133 in adult mouse tissues. (**d**) Northern blot analysis of the expression of miR-1 and miR-133 in embryonic day 13.5 (E13.5) and 16.5 (E16.5) mouse tissues. (**e**) Northern blot analysis of the expression of miR-1 and miR-133 in neonatal mouse tissues. Same amount of total RNAs from adult heart and skeletal muscle were loaded into blots to serve as a comparison to embryonic and neonate RNA (**d, e**).

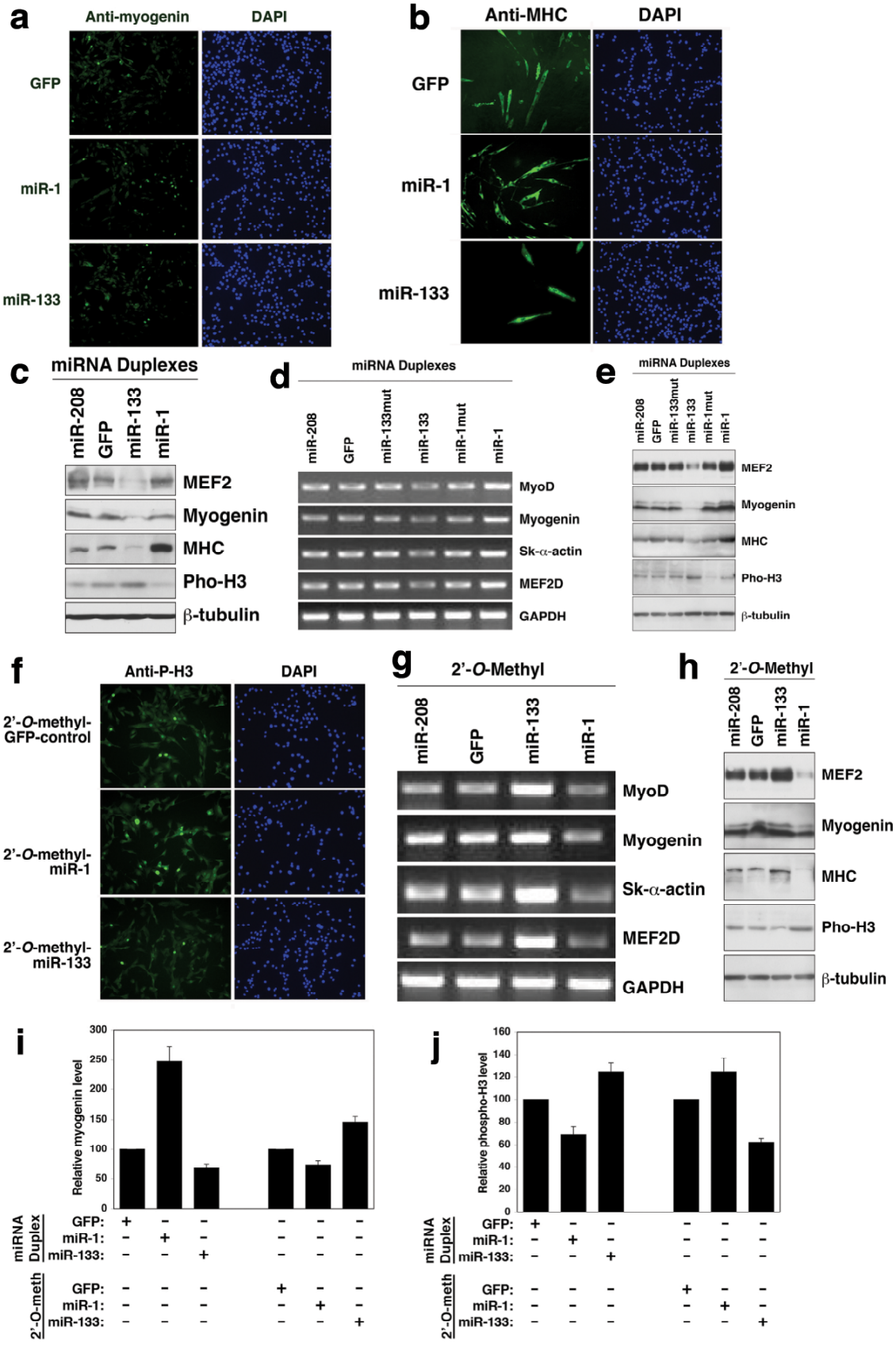

a
Anti-myogenin
DAPI
GFP
miR-1
miR-133
b
Anti-MHC
DAPI
GFP
miR-1
miR-133
c
miRNA Duplexes
miR-208
GFP
miR-133
miR-1
MEF2
Myogenin
MHC
Pho-H3
β-tubulin
d
miRNA Duplexes
miR-208
GFP
miR-133mut
miR-133
miR-1mut
miR-1
MyoD
Myogenin
Sk-α-actin
MEF2D
GAPDH
e
miRNA Duplexes
miR-208
GFP
miR-133mut
miR-133
miR-1mut
miR-1
MEF2
Myogenin
MHC
Pho-H3
β-tubulin
f
Anti-P-H3
DAPI
2'-O-methyl-GFP-control
2'-O-methyl-miR-1
2'-O-methyl-miR-133
g
2'-O-Methyl
miR-208
GFP
miR-133
miR-1
MyoD
Myogenin
Sk-α-actin
MEF2D
GAPDH
h
2'-O-Methyl
miR-208
GFP
miR-133
miR-1
MEF2
Myogenin
MHC
Pho-H3
β-tubulin
i
Relative myogenin level
0
50
100
150
200
250
300
miRNA Duplex
GFP: + − − − − −
miR-1: − + − − − −
miR-133: − − + − − −
2'-O-meth
GFP: − − − + − −
miR-1: − − − − + −
miR-133: − − − − − +
j
Relative phospho-H3 level
0
20
40
60
80
100
120
140
160
miRNA Duplex
GFP: + − − − − −
miR-1: − + − − − −
miR-133: − − + − − −
2'-O-meth
GFP: − − − + − −
miR-1: − − − − + −
miR-133: − − − − − +

Figure 5.2. Regulation of myoblast proliferation and differentiation by miR-1 and miR-133. C2C12 myoblasts cultured in growth medium (GM) were electroporated with double-stranded miRNA duplexes for miR-1, miR-133, and GFP as a control. (**a** to **e**) Cells were continuously cultured in GM for 24 hr after transfection, then transferred to differentiation medium (DM) for **(a)** 12 hr before immunostaining for myogenin or **(b)** 36 hr before immunostaining for MHC. C2C12 myoblasts cultured in GM were electroporated with double-stranded miRNA duplexes for miR-1, miR-133 (or their mutants as indicated), or miR-208 and GFP as controls and cultured for 24 hr before: **(c)** Western blotting using indicated antibodies; (**d**) cells were transferred to DM for 24 hr and RT-PCR for the indicated genes were performed; (**e**) cells were transferred to DM for 24 hr and Western blotting using the indicated antibodies. (**f** to **h**) C2C12 myoblasts cultured in GM were electroporated with 2'-*O*-methyl antisense oligonucleotide inhibitors for miR-1, miR-133 or miR-208 and GFP as cotrols. Cells were cultured in GM for 24 hr after transfection then transferred into DM for: **(f)** 12 hr before immunostaining for phospho-histone H3; (**g**) 24 hr before performing RT-PCR for the indicated genes; (**h**) 24 hr before Western blotting using indicated antibodies. (**i** and **j**) C2C12 myoblasts cultured in GM were electroporated with either the miRNA duplexes or 2'-*O*-methyl antisense oligonucleotide inhibitors as indicated. Cells were cultured in GM for 24 hr after transfection, then transfer into DM for 12 hr before immunostaining for myogenin (**i**) or phospho-histone H3 (**j**). Positive stained cells were counted and data are presented as the expression level relative to a GFP control (100%).

	DM (8 hr)		DM (12 hr)				DM (24 hr)					
Treatment	**Myogenin positive cells**	**Relative to control**	**Myogenin positive cells**	**Relative to control**	**Phospho-H3 positive cells**	**Relative to control**	**Myogenin positive cells**	**Relative to control**	**Phospho-H3 positive cells**	**Relative to control**	**MHC positive cells**	**Relative to control**
GFP	172	100%	93	100%	135	100%	118	100%	137	100%	22	100%
miR-1	206	121%	230	247.30%	93	68.90%	251	212.70%	76	55.50%	56	254.50%
miR-133	89	51.70%	68	73.10%	168	124.40%	93	78.80%	201	146.70%	12	54.50%
2'-O-methyl-GFP	146	100%	145	100%	172	100%	348	100%	207	100%	22	100%
2'-O-methyl-miR-1	120	82.20%	98	67.60%	214	124.40%	299	85.90%	283	136.70%	18	81.80%
2'-O-methyl-miR-133	205	140.40%	211	145.50%	107	62.20%	498	143.10%	191	92.30%	44	200%

Table 5.1. Effect on myogenic proliferation and differentiation by miR-1 and miR-133 overexpression and knock down. C2C12 myoblasts cultured in growth medium (GM) were electroporated with double-stranded miRNA duplex or 2'-*O*-methyl antisense oligos for miR-1, miR-133 or GFP as a negative control. 36 hr later, GM was replaced with differentiation medium (DM) for 8, 12 and 24 hr and cells were fixed for immunohistochemistry analysis using antibodies against myogenin, phospho-histone H3 and Myosin heavy chain (MHC). Positive cells were counted out of 5000 DAPI staining cells from a randomly chosen field. Assays were performed three times independently with comparable results.

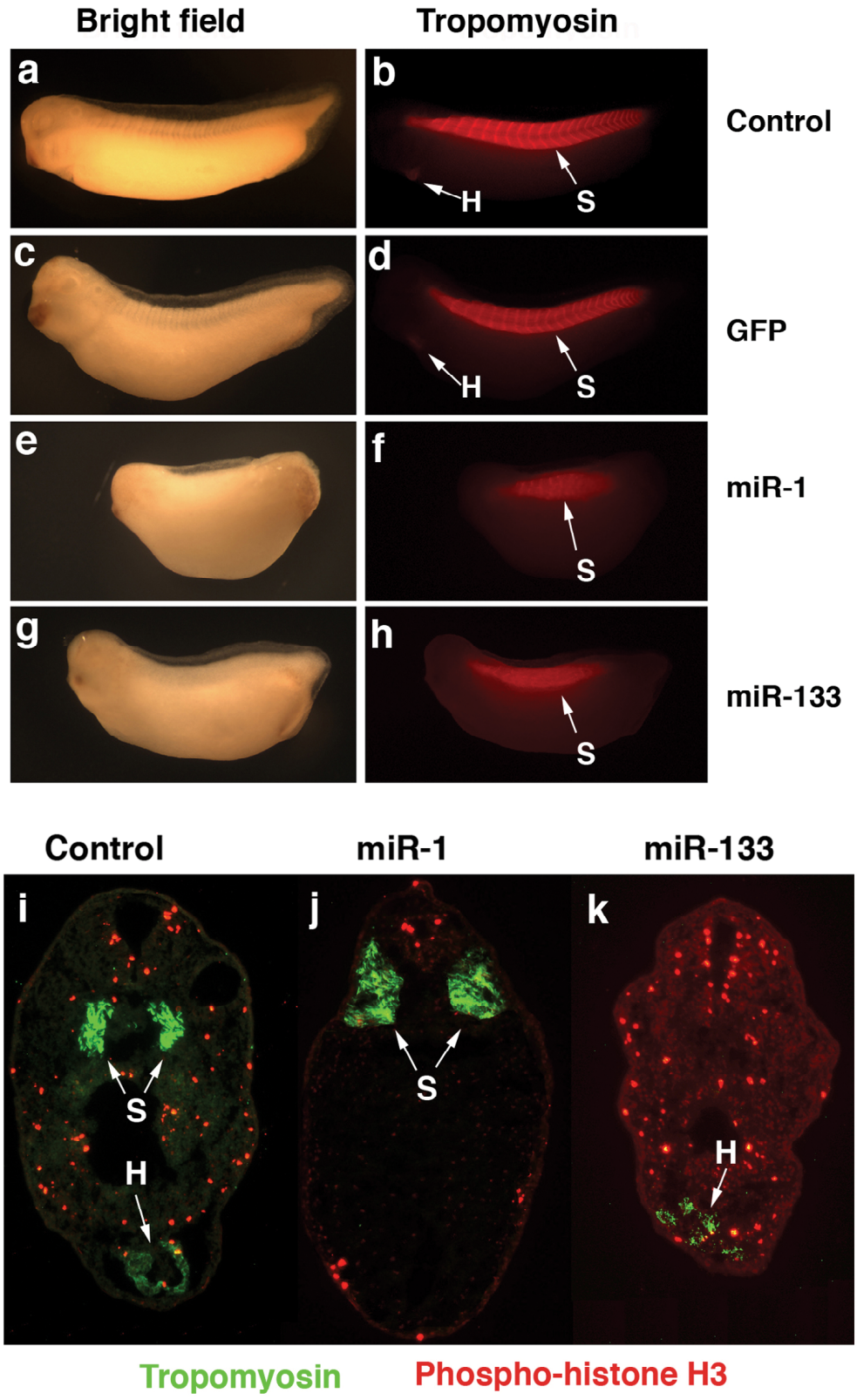

Bright field
Tropomyosin
a
b
Control
H
S
c
d
GFP
H
S
e
f
miR-1
S
g
h
miR-133
S
Control
miR-1
miR-133
i
j
k
S
H
S
H
Tropomyosin
Phospho-histone H3

Figure 5.3. Control of cardiac and skeletal muscle development by miR-1 and miR-133 *in vivo*. Top panels: *Xenopus* embryos derived from uninjected (**a** and **b**), GFP RNA control-injected (**c** and **d**), miR-1-injected (**e** and **f**), or miR-133-injected (**g** and **h**) embryos stained with anti-tropomyosin and shown at stage 32 under brightfield (**a, c, e, g**) or fluorescence (**b, d, f, h**). Note the lack of staining for heart tissue (H, arrows) and disruption of segmented somites (S, arrows) in **f** and **h**. Bottom panels: Transverse sections of *Xenopus* embryos corresponding to the position of the heart at stage 32 from uninjected (**i**), miR-1 injected (**j**), or miR-133 injected (k) embryos stained with anti-tropomyosin to visualize somites (S, arrows) and cardiac tissue (H, arrows), and anti-phospho-histone H3 (red) to visualize cells in S phase. Each set of injections was conducted at least twice independently, and the phenotype was observed in at least 90% of a minimum of 50 embryos scored by whole mount immunostaining.

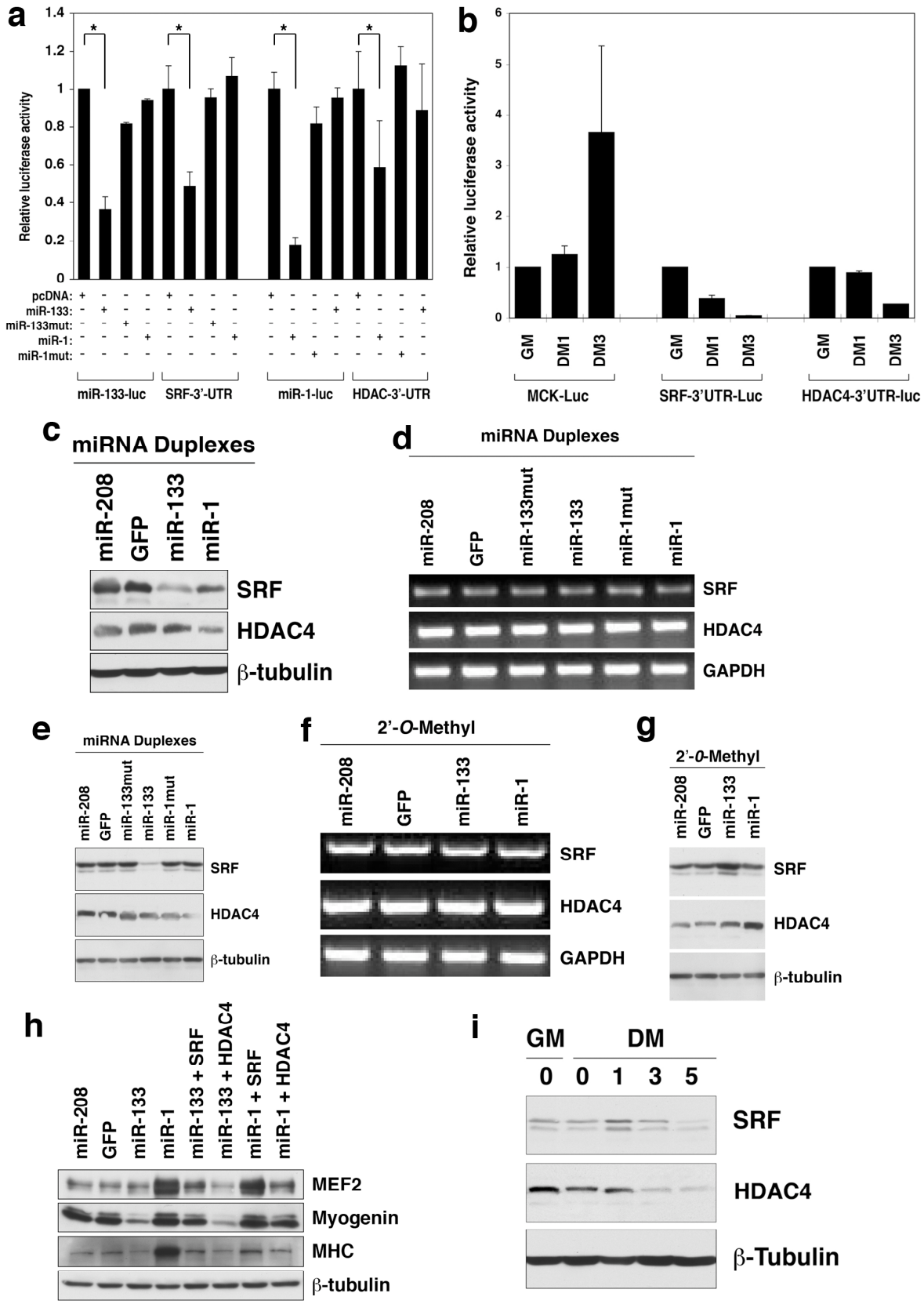

a
Relative luciferase activity
pcDNA:
miR-133:
miR-133mut:
miR-1:
miR-1mut:
miR-133-luc
SRF-3'-UTR
miR-1-luc
HDAC-3'-UTR
b
Relative luciferase activity
GM
DM1
DM3
MCK-Luc
SRF-3'UTR-Luc
HDAC4-3'UTR-luc
c
miRNA Duplexes
miR-208
GFP
miR-133
miR-1
SRF
HDAC4
β-tubulin
d
miRNA Duplexes
miR-208
GFP
miR-133mut
miR-133
miR-1mut
miR-1
SRF
HDAC4
GAPDH
e
miRNA Duplexes
miR-208
GFP
miR-133mut
miR-133
miR-1mut
miR-1
SRF
HDAC4
β-tubulin
f
2'-O-Methyl
miR-208
GFP
miR-133
miR-1
SRF
HDAC4
GAPDH
g
2'-O-Methyl
miR-208
GFP
miR-133
miR-1
SRF
HDAC4
β-tubulin
h
miR-208
GFP
miR-133
miR-1
miR-133 + SRF
miR-133 + HDAC4
miR-1 + SRF
miR-1 + HDAC4
MEF2
Myogenin
MHC
β-tubulin
i
GM
DM
0
0
1
3
5
SRF
HDAC4
β-Tubulin

Figure 5.4. Identification of miR-1 and miR-133 target genes in skeletal muscle. (**a**) Repression of SRF and HDAC4 3'UTRs by miR-133 and miR-1. Luciferase reporters containing either miR-133 complementary sites from mouse SRF 3' UTR (SRF-3'-UTR), miR-1 complementary sites from mouse HDAC4 3' UTR (HDAC4-3'-UTR) or the perfect antisense sequences of miR-133 (miR-133-luc) or miR-1 (miR-1-luc) were co-transfected with the indicated miRNA expression vectors or their mutants. Luciferase activity was determined 48 hr after transfection. Data represent the mean $\pm$ s.d. from at least three independent experiments in duplicate (* $P < 0.05$). (**b**) SRF-3'-UTR, HDAC4-3'-UTR, and MCK-luc luciferase reporters were transfected into C2C12 myoblasts. Cells were maintained in GM for 24 hr (GM) or transferred into DM for 1 day (DM1) or 3 days (DM3) before luciferase activity was determined. (**c** to **e**) C2C12 myoblasts cultured in GM were electroporated with indicated double-stranded miRNA duplexes (or their mutants), or miR-208 and GFP as controls. Cells were cultured in GM for 24 hr after transfection before: (**c**) Western blotting using anti-SRF and anti-HDAC4 antibodies; (**d**) cells were transferred into DM for 24 hr and RT-PCR for the indicated genes performed; (**e**) cells were transferred into DM for 24 hr and Western blotting using indicated antibodies. C2C12 myoblasts cultured in GM were electroporated with indicated 2'-*O*-methyl antisense oligonucleotide inhibitors. (**f** and **g**) Cells were cultured in GM for 24 hr after transfection, then transferred into DM for 24 hr before: (**f)** RT-PCR for the indicated genes performed; (**g)** Western blotting using indicated antibodies. (**h**) C2C12 myoblasts cultured in GM were electroporated with indicated double-stranded miRNA duplexes or/and expression plasmids for SRF or HDAC4, as indicated. Cells were cultured in GM for 24 hr after transfection. Western blotting

performed 24 hr after transfer into DM using indicated antibodies. (**i**) C2C12 myoblasts were cultured in GM or DM for 0, 1, 3 or 5 days. Western blotting using indicated antibodies.

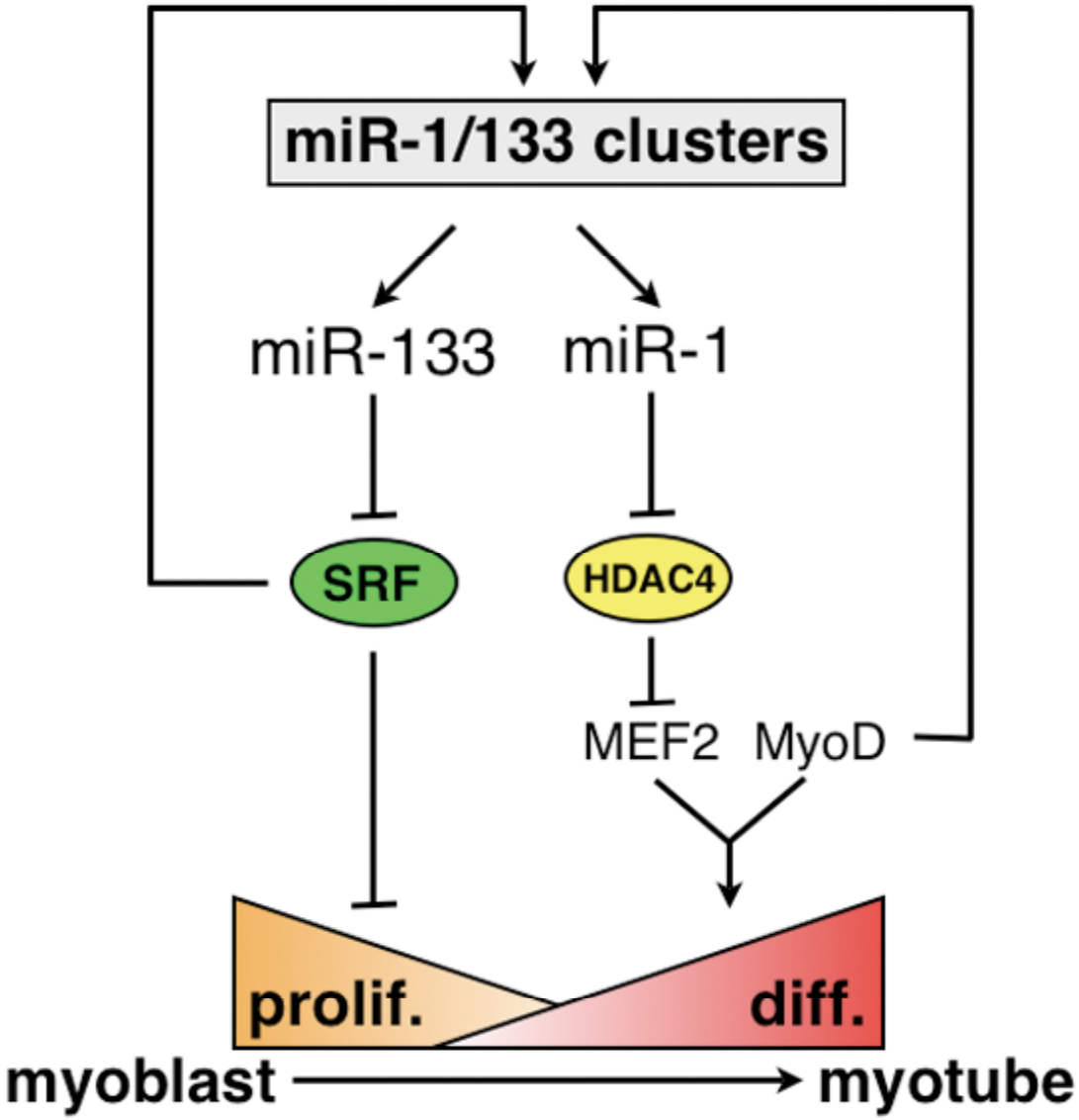


Figure 5.5. Model for miR-1 and miR-133-mediated regulation of skeletal muscle proliferation and differentiation.

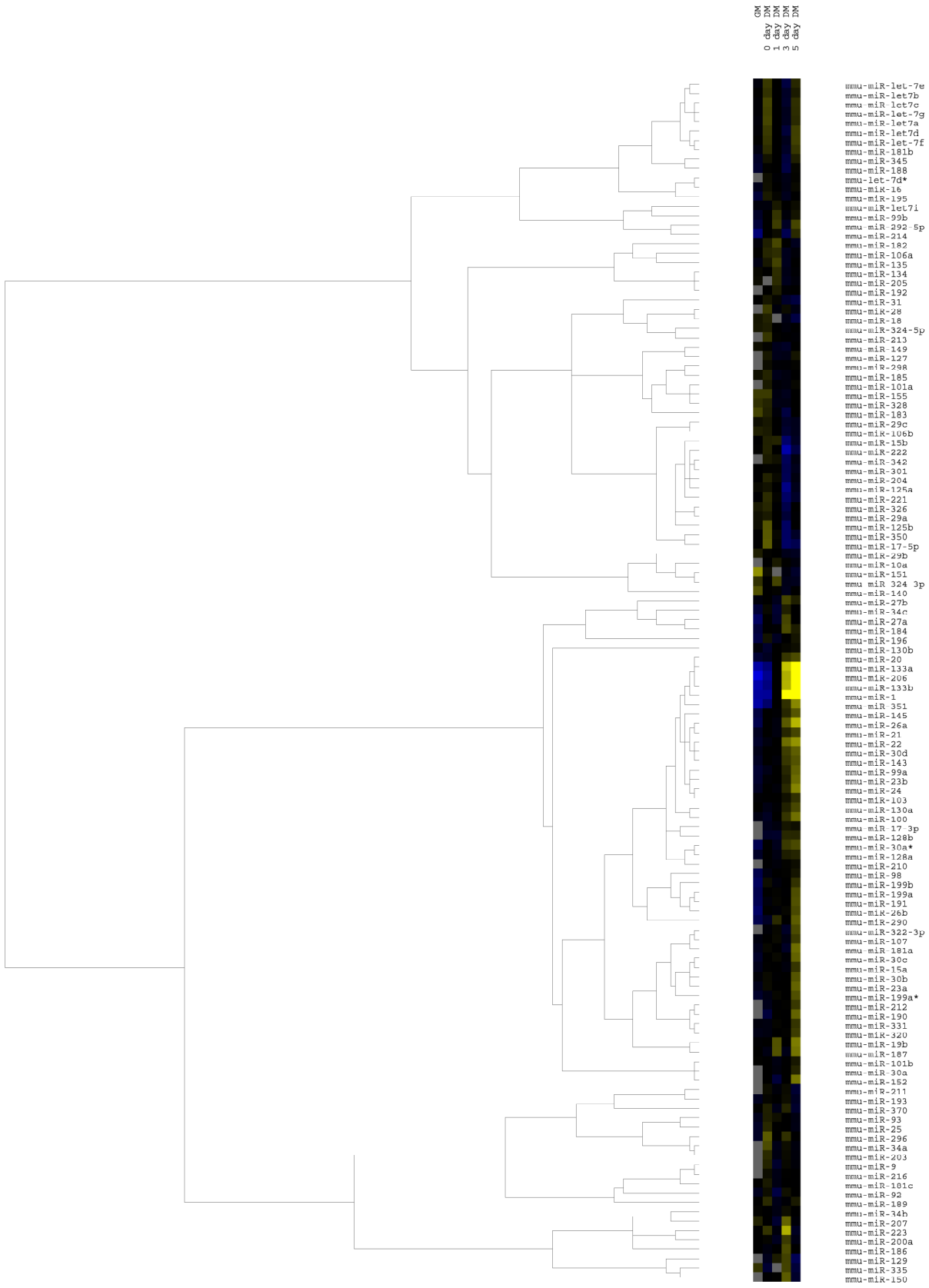

GM
0 day DM
1 day DM
3 day DM
5 day DM
mmu-miR-let-7e
mmu-miR-let7b
mmu-miR-let7c
mmu-miR-let-7g
mmu-miR-let7a
mmu-miR-let7d
mmu-miR-let-7f
mmu-miR-181b
mmu-miR-345
mmu-miR-188
mmu-let-7d*
mmu-miR-16
mmu-miR-195
mmu-miR-let7i
mmu-miR-99b
mmu-miR-292-5p
mmu-miR-214
mmu-miR-182
mmu-miR-106a
mmu-miR-135
mmu-miR-134
mmu-miR-205
mmu-miR-192
mmu-miR-31
mmu-miR-28
mmu-miR-18
mmu-miR-324-5p
mmu-miR-213
mmu-miR-149
mmu-miR-127
mmu-miR-298
mmu-miR-185
mmu-miR-101a
mmu-miR-155
mmu-miR-328
mmu-miR-183
mmu-miR-29c
mmu-miR-106b
mmu-miR-15b
mmu-miR-222
mmu-miR-342
mmu-miR-301
mmu-miR-204
mmu-miR-125a
mmu-miR-221
mmu-miR-326
mmu-miR-29a
mmu-miR-125b
mmu-miR-350
mmu-miR-17-5p
mmu-miR-29b
mmu-miR-10a
mmu-miR-151
mmu-miR-324-3p
mmu-miR-140
mmu-miR-27b
mmu-miR-34c
mmu-miR-27a
mmu-miR-184
mmu-miR-196
mmu-miR-130b
mmu-miR-20
mmu-miR-133a
mmu-miR-206
mmu-miR-133b
mmu-miR-1
mmu-miR-351
mmu-miR-145
mmu-miR-26a
mmu-miR-21
mmu-miR-22
mmu-miR-30d
mmu-miR-143
mmu-miR-99a
mmu-miR-23b
mmu-miR-24
mmu-miR-103
mmu-miR-130a
mmu-miR-100
mmu-miR-17-3p
mmu-miR-128b
mmu-miR-30a*
mmu-miR-128a
mmu-miR-210
mmu-miR-98
mmu-miR-199b
mmu-miR-199a
mmu-miR-191
mmu-miR-26b
mmu-miR-290
mmu-miR-322-3p
mmu-miR-107
mmu-miR-181a
mmu-miR-30c
mmu-miR-15a
mmu-miR-30b
mmu-miR-23a
mmu-miR-199a*
mmu-miR-212
mmu-miR-190
mmu-miR-331
mmu-miR-320
mmu-miR-19b
mmu-miR-187
mmu-miR-101b
mmu-miR-30a
mmu-miR-152
mmu-miR-211
mmu-miR-193
mmu-miR-370
mmu-miR-93
mmu-miR-25
mmu-miR-296
mmu-miR-34a
mmu-miR-203
mmu-miR-9
mmu-miR-216
mmu-miR-101c
mmu-miR-92
mmu-miR-189
mmu-miR-34b
mmu-miR-207
mmu-miR-223
mmu-miR-200a
mmu-miR-186
mmu-miR-129
mmu-miR-335
mmu-miR-150

Figure S.5.1. miRNA array analysis of C2C12 cells. miRNA array expression data from C2C12 myoblasts cultured in growth medium (GM) or in differentiation medium (DM) for 0, 1, 3 and 5 days, respectively. Normalized log (base 2) data was hierarchically clustered by gene and is plotted as a heat map. The range of signal was from –4 fold to +4 fold. Yellow denotes high expression and blue denotes low expression, relative to the median.

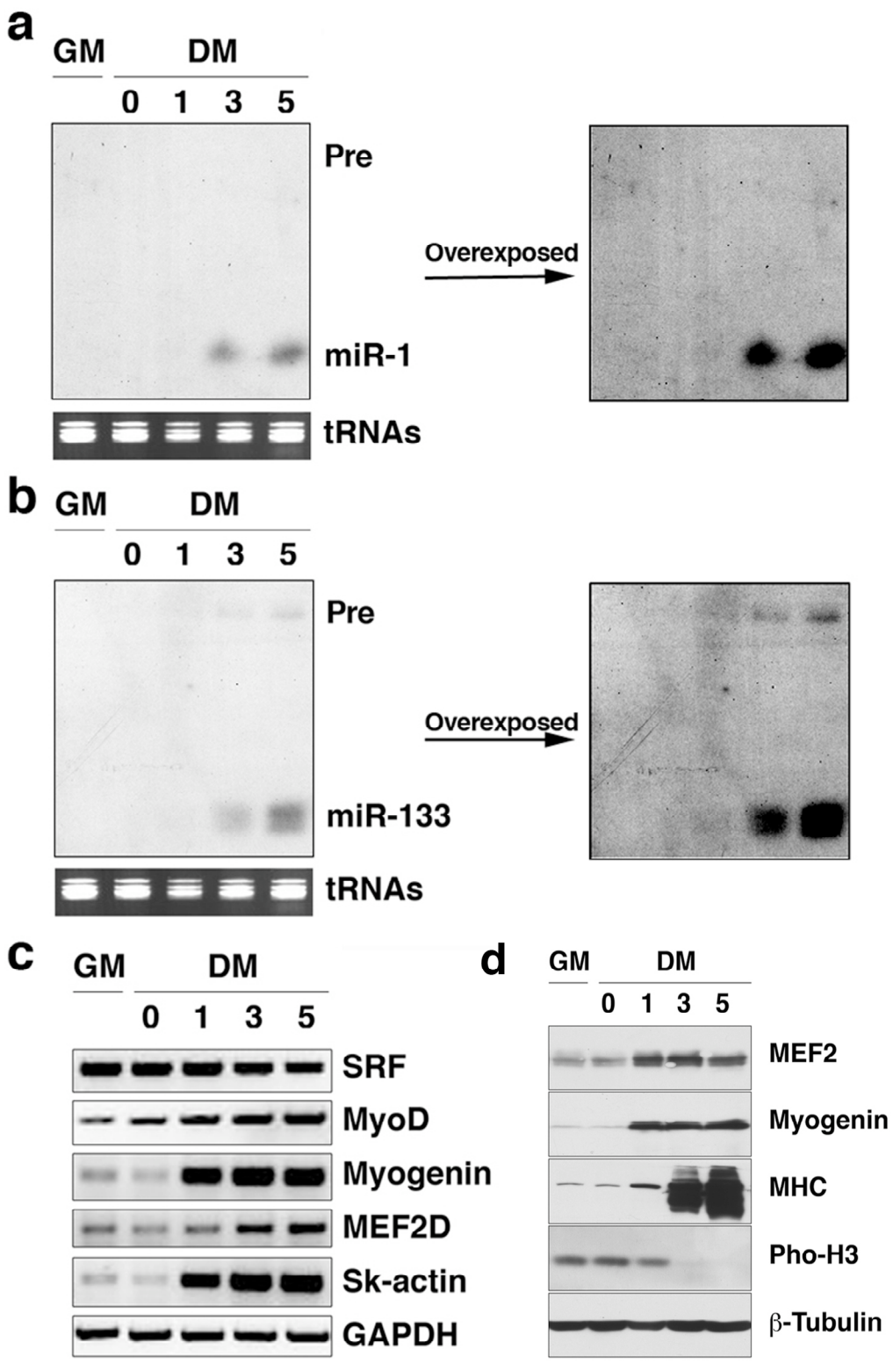

a
GM
DM
0 1 3 5
Pre
miR-1
tRNAs
Overexposed
b
GM
DM
0 1 3 5
Pre
miR-133
tRNAs
Overexposed
c
GM
DM
0 1 3 5
SRF
MyoD
Myogenin
MEF2D
Sk-actin
GAPDH
d
GM
DM
0 1 3 5
MEF2
Myogenin
MHC
Pho-H3
β-Tubulin

Figure S.5.2. Expression of miR-1, miR-133 and skeletal muscle differentiation marker genes in C2C12 cells. (**a** and **b**) Northern blot analysis of the expression of miR-1 (**a**) and miR-133 (**b**) using total RNA isolated from C2C12 myoblasts cultured in GM or in differentiation medium (DM) for 0, 1, 3 and 5 days, respectively. Both mature miRNAs and their precursors (Pre) are indicated. tRNAs were used as a loading control. (**c**) Semi-quantative RT-PCR analysis of skeletal muscle differentiation marker genes. GAPDH was used as a control for equal loading. (**d**) Expression of skeletal muscle differentiation markers. C2C12 myoblasts were cultured in growth medium (GM) or in differentiation medium (DM) for 0, 1, 3 and 5 days, and Western blots performed with cell extracts using the indicated antibodies. β-tubulin serves as a loading control.

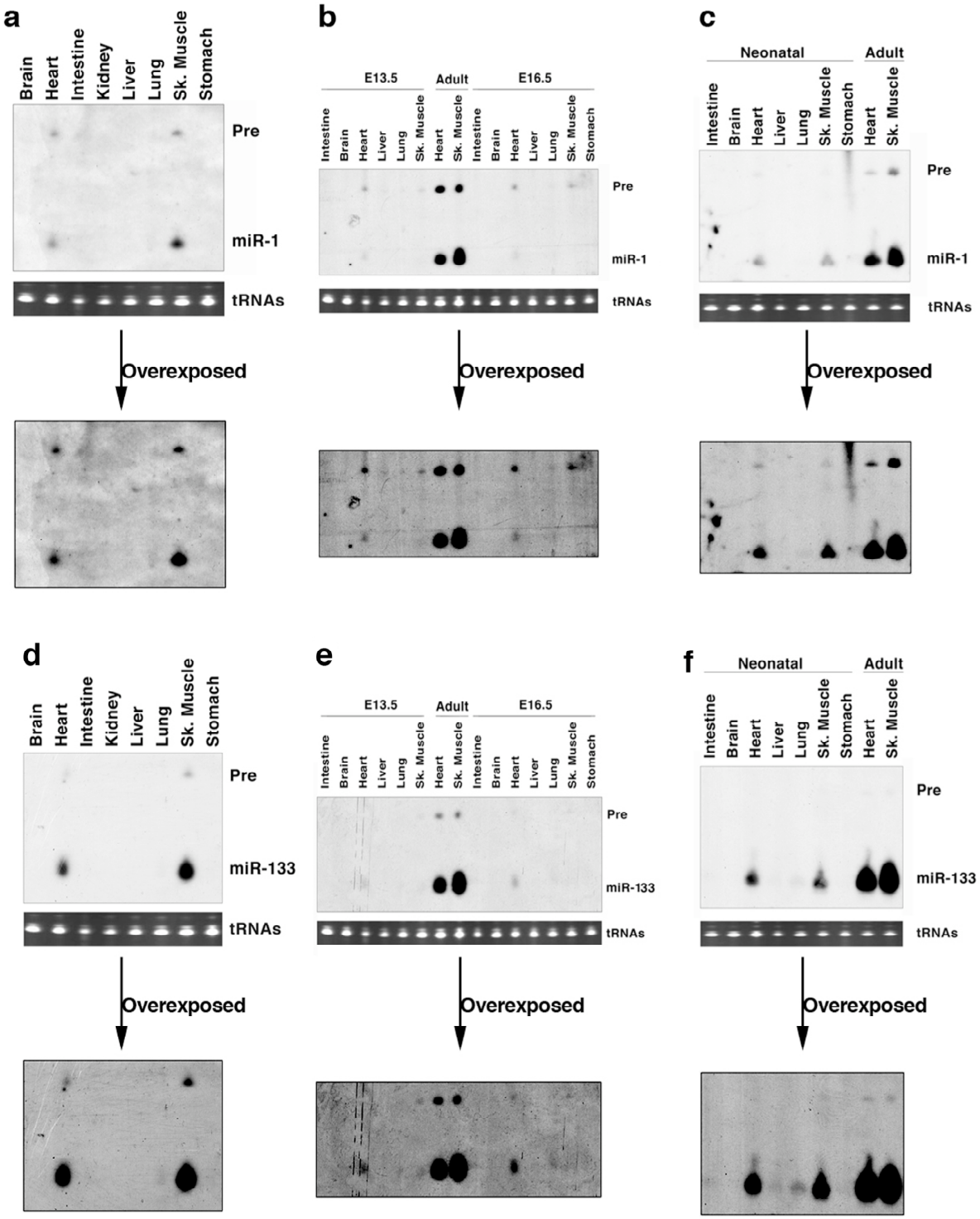

a
Brain
Heart
Intestine
Kidney
Liver
Lung
Sk. Muscle
Stomach
Pre
miR-1
tRNAs
Overexposed
b
E13.5
Adult
E16.5
Intestine
Brain
Heart
Liver
Lung
Sk. Muscle
Heart
Sk. Muscle
Intestine
Brain
Heart
Liver
Lung
Sk. Muscle
Stomach
Pre
miR-1
tRNAs
Overexposed
c
Neonatal
Adult
Intestine
Brain
Heart
Liver
Lung
Sk. Muscle
Stomach
Heart
Sk. Muscle
Pre
miR-1
tRNAs
Overexposed
d
Brain
Heart
Intestine
Kidney
Liver
Lung
Sk. Muscle
Stomach
Pre
miR-133
tRNAs
Overexposed
e
E13.5
Adult
E16.5
Intestine
Brain
Heart
Liver
Lung
Sk. Muscle
Heart
Sk. Muscle
Intestine
Brain
Heart
Liver
Lung
Sk. Muscle
Stomach
Pre
miR-133
tRNAs
Overexposed
f
Neonatal
Adult
Intestine
Brain
Heart
Liver
Lung
Sk. Muscle
Stomach
Heart
Sk. Muscle
Pre
miR-133
tRNAs
Overexposed

Figure S.5.3. Expression of miR-1 and miR-133 in cardiac and skeletal muscle in adult mice and throughout development. Northern blot analysis of the expression of miR-1 (**a**) and miR-133 (**d**) in adult mouse tissues. Northern blot analysis of the expression of miR-1 (**b**) and miR-133 (**e**) in embryonic day 13.5 (E13.5) and 16.5 (E16.5) mouse tissues. Same amount of total RNA from adult heart and skeletal muscle was also loaded in the blot to serve as a comparison. Northern blot analysis of the expression of miR-1 (**c**) and miR-133 (**f**) in neonatal mouse tissues. Same amount of total RNA from adult heart and skeletal muscle was also loaded in the blot to serve as a comparison. Both mature miRNAs and their precursors (Pre) are indicated. tRNAs were used as a loading control.

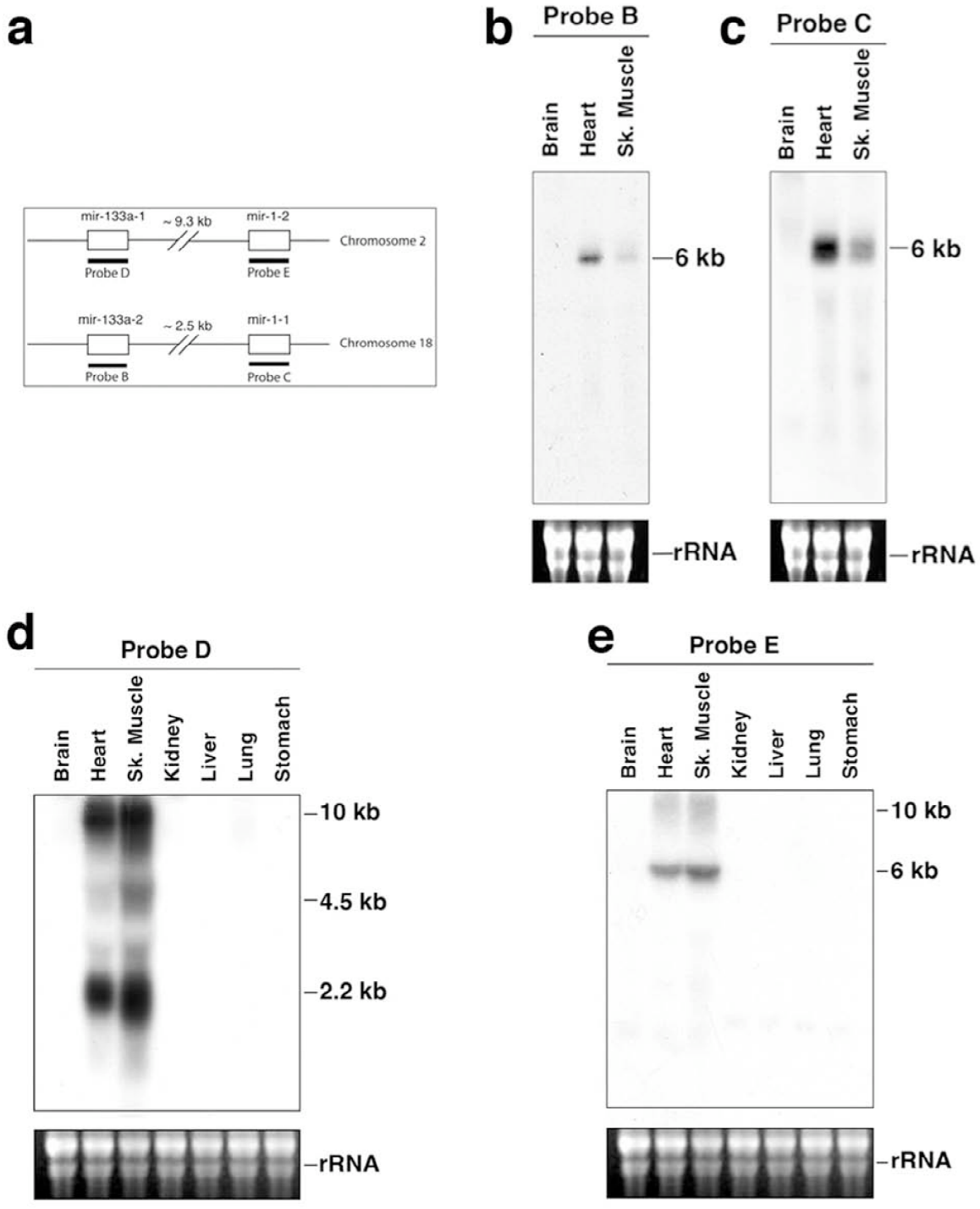

a
mir-133a-1
~ 9.3 kb
mir-1-2
Chromosome 2
Probe D
Probe E
mir-133a-2
~ 2.5 kb
mir-1-1
Chromosome 18
Probe B
Probe C
b
Probe B
Brain
Heart
Sk. Muscle
6 kb
rRNA
c
Probe C
Brain
Heart
Sk. Muscle
6 kb
rRNA
d
Probe D
Brain
Heart
Sk. Muscle
Kidney
Liver
Lung
Stomach
10 kb
4.5 kb
2.2 kb
rRNA
e
Probe E
Brain
Heart
Sk. Muscle
Kidney
Liver
Lung
Stomach
10 kb
6 kb
rRNA

Figure S.5.4. Expression of miR-1 and miR-133 primary transcripts in cardiac and skeletal muscle. (**a**) Diagram showing miR-1 and miR-133 genes clustered on mouse chromosomes 2 and 18. Probes used for Northern blots in **b** through **e** are noted. (**b** - **e**) Northern blot analysis of the expression of primary transcripts for miR-1 (**c**, **e**) and miR-133 (**b**, **d**) from chromosome 2 (**d**, **e**) and chromosome 18 (**b**, **c**). 20 μg of total RNA from the indicated adult mouse tissues was used.

Chromosome 2 miR-1/133 cluster enhancer sequences

```
   1 GAGCAAGTTT CACTAGGGCC ACACAGTATC ATTGAGCACT GAGCGTGGAA GGAGACAGAT
  61 GGGCCACGTT TCTCCTCCCT CTTTCTAGCC TTCCTTCTCC CTCCCTTTTC TTATACATTA
 121 TATCCTGGCG GCAGTTTTCC CTCCCTCCAC TCCTCCCAGT TCCTTCCCCA CTTCCATTCT
 181 CCCCCAGATC CATTCCTCTT ATGCCCCCCC CCCCCAAGAG CAGGCTTTCC ATGAATACCC
 241 ACCAAACATG GCATAACAAG TTACAATAAG ATCAGGAACA AACCCTCATA TCAAGGCTGG
 301 ATGAGGCAAC CCAACAGGAG GAAAAGGGCC CCAAGAGCAG GCAAAAAACT CCCACTGTTG
 361 TGTCTTCTGC TAGAACACAA AGCTACACAA CTATAATGTA TATGCAGAGG ACCCAGCTCA
 421 GTCTCATCAG GGTCCGTGTT TGTTGCTACA GTTTCTGTGA ACCTCTGTGG GCTCTGCTTA
 481 GTTGGTTCTG TGGGTTGTGT TCTTGTGGCA TCCTCAACTC CTCTGGCTCC TACAATCTTT
 541 CCTCCCATCT TCTTTGGAGT TCCCCTGGCC ATGCCTGATG TTTGGTTGGC TTGGCTGTGT
 601 GGGCCTCTGC ATTTATTTCC GTCAGTTGCT GGAAAGCATC CCTCTGTTGA CAGTTGGTCC
 661 ATGCACTGAT CTATGAGGAT AGCAGAGTAT CACTAGGAAT TACTTTATTG TCTTTTTTGC
 721 CAGTCGTTTT TGGTTCTCTC CCGAGTCTCT GGGCTGTCCA GTCTCTGGTT CCTGGCCTTC
 781 CAGACACTGT CAGTTGTGGG TTCCCTTTTG TGGTGTTGGC CTCAACTTGG CCAGTCATTG
 841 GTTGGCCATT CCCACAAGTT CTACACCACC ATTACCCTAG CATGTCTCGC AGGCAGGACA
 901 GATTGTACGC GGAAGGATTT ATGGCTGGGT TATGTCTCAG TCCCAGGGCT GGAAGCCTTG
 961 CCTGGTTAAA GAAGACAGCT AGTTCTGACT CAGTATTCCC TGTTACTAGC AGAATTCACT
1021 AGGATTACCC TCACCCTCAG GGCATTTCCA CAGCACTAGG GTTCTGCATT GCCTCTCCAA
1081 TACCCCCTCC AATTCCAGTC GCCTTTCCCA GAACTCTCCT CCCCCAGCCT GATCCCTATT
1141 GTTCCCACCC CCATCCACCC CCAGTCCACC TACAAAGCTC TTTCCCCTTC CCAAGAAGAT
1201 CCATGAGTTT CTCTGTGTCT GTGGATTGGA GTATGATCTG TGGATTTAGC AGCTAATGTT
1261 CACTTACCAG TGAGTACACA CCGTTTGTCT TTTGGGTCTG GGTTACCTCA CTCAGGGTGG
1321 ATTTTGGATT TTTTTTTTTG AGTGCTATCC ATTTGTCTGC AAATGTCATG ATGTCATTTT
1381 TTTTAACAGC TGAGGAATTC TCTCAGAACC ACATTTTCTT TATCCATTAT TCAGATTGTT
                   MEF2 site
1441 CCCAGTTTCT GGCTATTATA AGGCTGCTAT GAACATGGTT GAACAAGTGT CCTTGTGGTA
                 tcggTAcA ca
                (MEF mutation)
1501 GGATGTGGCA TCCTTTGGGT ATATGCCTAG GAGTGGTATC GATGGGTCTC GAGGTAGATC
1561 AATTCCCGAT TTTCTGAGAA ACTGCCATAT CTGTTTCCAA AGTGGCTGTG TAAGTTTGCG
1621 CTCCCACCAG CAACGGAGGA GTGTTCTCCT TACTCCTCCC ACATTATCAA CAGTGTGAGC
1681 TGTCACTTGT GTTTTTGATC TTAGCCTTTC TGACAGGTGT AAGATGGAAT CTCAAAGTAG
1741 CTTTGATTTG CATTTCCCTG CTGGCTAAGG ATGTTGAACA TTTCTTTAAG TGTTTCTCAG
1801 CCATTTGAGA TTTATCCATT GAGAATTCTG TTTAGATCTG AACTCCACCT TCTAATTGGA
1861 TTATTTGGTT TTTAAAATAT CCACTTTCTC GAGTTCTTAA TGGGTTTTGG ATATTAGCCC
1921 TCTGTCAAAT GTGGAGTTGG TGAAGATCTT TTCCCATTCG GTAGGTTTTG TCCTATTGAC
                                                 CArG Box
1981 AGTGTCCTTT GCTTCACAGA AGCTTTTCAG TTTCATGAGG TCCCATTTAT TGATTGTTGA
                                                   atcTTggT ct
                                                 (CArG Box mutation)
2041 TCTTAGTGCC TGTGCTATTG ATGTCTATTC AGGAAGTTGT CTCCTGTGCC AATGCGTTCA
2101 AGGCTATTTC CCACTTTCTC TTCTATTAGG TTCAGTGTAT CTCATTTTAT GTTGAGGTCT
2161 TTGATCCACT TAGAGTTGAG TTTTGTGCAG AGTGATAGAT ATGGATCTAT TTGCATTCTT
2221 CTACATGCAG ATATCCAGTA AGACCAGCAT CATTTATTGC GGATGCTTTT TAAATTTTTT
2281 CGCTTGTGTA TTTCTGGCTT CTTTATAAAA ATCAGGTGTT CACTGATTTC ATTGATCAGC
2341 CAATGCTTTT CTGCCGATAC CATGTGGTTT TATTGCTATA GCTCTGAGGT ACAGCTTGAG
2401 TCAGGGATGT GATGCCCCTG GACGTCCTTT TATTGTACAG GAGTATCCTA GGTTTAGCTA
2461 TCCTAGGTTT TTTGGTTTTC CACATGGAGT TAAGTATTGT CCTTTCAAGG TCTATAGAGA
2521 ATTGCATTGG GATTTTGGTG GAGATTGTAT TGCATTTGTA GATTTGGTAG GGTGGCCATT
2581 TTTACTATGG TAATCCTACC
```

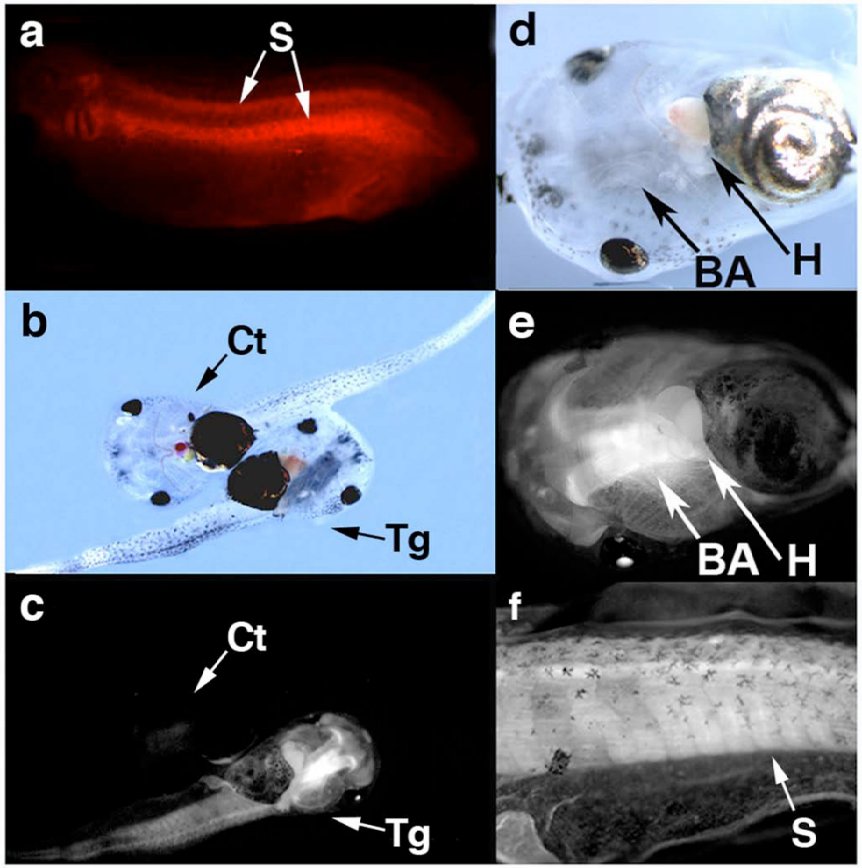


Figure S.5.5. A miR-1 and miR-133 enhancer directs reporter gene expression in cardiac and skeletal muscle. Genomic DNA sequences of miR-1/133 enhancer from mouse chromosome 2. A putative MEF2 site and CArG box are marked out, and mutations introduced into these sites are indicated (upper panel). (**a**) *Xenopus laevis* transgenic for mouse miR-1 and miR-133 genomic sequence linked to dsRed shows somite (S, arrows) expression at stage 28. (**b**) Transgenic (Tg) *Xenopus laevis* carrying a miR-1 and miR-133-containing transgene at stage 46 (lower embryo) and negative control (non-transgenic, Ct, upper embryo) under bright field. (**c**) Same embryos as b under fluorescence. (**d**) High power magnification of transgenic embryo in panel g under bright field and under fluorescence (**e**) showing expression of the trangene in the heart (H, arrows) and branchial arches (BA, arrows). (**f**) High magnification of stage 46 transgenic embryo showing expression of the transgene in the somites (S, arrows).

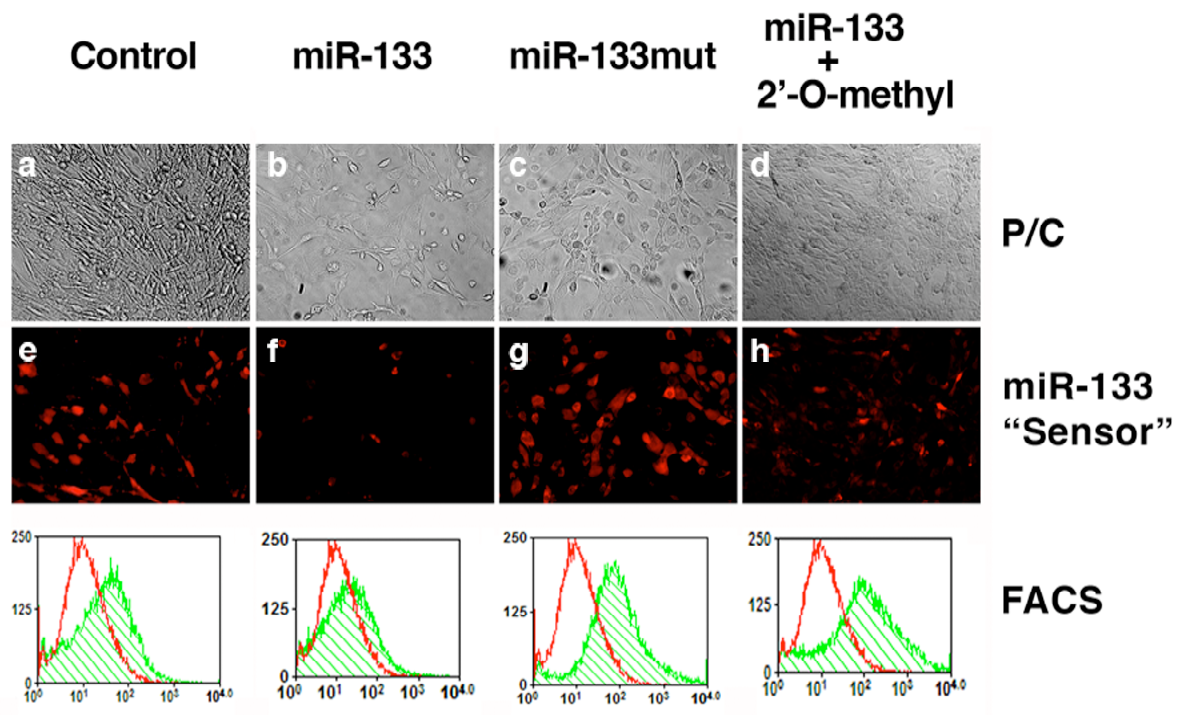


Figure S.5.6. Repression of a miR-133 sensor by miR-133 in C2C12 cells. C2C12 cells stably expressing the miR-133 sensor were transfected with expression vectors for GFP (control), wild –type miR-133 (miR-133), mutant miR-133 (miR-133mut) in which the "seed" sequence has been mutated, or a combination of miR-133 expression vector and 2'-O-methyl antisense oligos (miR-133 + 2'-O-methyl). Cells were transferred into differentiation medium for 12 hr and images were obtained using phasecontrast (P/C) (**a, b, c, d**) or fluorescence to show expression of the dsRed reporter gene (**e, f, g, h**). Cells from each condition were harvested and the expression of the dsRed reporter gene was quantified using FACS analysis (lower panels). Red line denotes autofluorescein of the cell and green line indicates the ds-Red expression.

Figure S.5.7. Sequences of the miR-1 and miR-133 target sites in the 3' UTR of HDAC4 and SRF genes. Upper panel: HDAC4 3' UTR sequences from conserved vertebrate species and their alignment with miR-1 and miR-206. Lower panel: SRF 3' UTR sequences from human and rat and their alignment with miR-133. Conserved nucleotide sequences are listed in upper case.

Table S.4.1 Sequences of oligonucleotides used in this study.

Name	Sequence
miR-1 probe	TACATACTTCTTTACATTCCA
miR-133 probe	ACAGCTGGTTGAAGGGGACCAA
miR-133a-1-up	CATGTGACCCCTCACACACA
miR-133a-1-down	ACAAGGGGAGCCTGGATCCC
miR-133a-2-up	GGACATATGCCTAAACACGTGA
miR-133a-2-down	GAAACATCTTTATCCAGTTT
miR-1-2-up	AGACTGAGACACAGGCGACACC
miR-1-2-down	TGCCGGTCCATCGGTCCATTGC
miR-1-1-up	CACTGGATCCATTACTCTTC
miR-1-1-down	TTGGAATGGGGCTGTTAGTA
miR-1mut-up	TGAACATTCAGTGCTATAAAGAAGTATGTATTTTGGGTAGGTA
miR-1mut-down	TACCTACCCAAAATACATACTTCTTTATAGCACTGAATGTTCA
miR-133mut-up	AATCGCCTCTTCAATGGATTTGTCAACCAGCTGTAGCTATGCATTGAT
miR-133mut-down	ATCAATGCATAGCTACAGCTGGTTGACAAATCCATTGAAGAGGCGATT
miR-1 duplex	UGGAAUGUAAAGAAGUAUGUA CAUACUUCUUUACAUUCCAUA
miR-1-mut duplex	UUAACCAUAAAGAAGUAUGUA CAUACUUCUUUAUGGUUAAUA
miR-133 duplex	UUGGUCCCCUUCAACCAGCUGU AGCUGGUUGAAGGGGACCAAAU
miR-133-mut duplex	UCAAGUAACUUCAACCAGCUGU AGCUGGUUGAAGUUACUUGAAU
miR-208 duplex	AUAAGACGAGCAAAAAGCUUGU AAGCUUUUUGCUCGUCUUAUAC
GFP duplex	AACUUCAGGGUCAGCUUGCCUU GGCAAGCUGACCCUGAAGUUGG
2'-*O*-methyl-miR-1	AAAUACAUACUUCUUUACAUUCCAUAGC
2'-*O*-methyl-miR-133	AGCUACAGCUGGUUGAAGGGGACCAAAUCCA
2'-*O*-methyl-miR-208	GACCAACAAGCUUUUUGCUCGUCUUAUACGUG
2'-*O*-methyl-GFP	AAGGCAAGCUGACCCUGAAGUU
HDAC4-UTR-up	CAGCACTGGTGATAGACTTGG
HDAC4-UTR-down	CTTAAGAATAAGTTCAATAAGAC
SRF-UTR-up	AGATATGGGGGCTTGTGCCC
SRF-UTR-down	CTGGGAGAAAGGGGGTAGAC
Myogenin F	TGGAGCTGTATGAGACATCCC
Myogenin R	TGGACAATGCTCAGGGGTCCC
MyoD F	GCAGGCTCTGCTGCGCGACC
MyoD R	TGCAGTCGATCTCTCAAAGCACC
Skeletal α-actin F	CAGAGCAAGCGAGGTATCC
Skeletal α-actin R	GTCCCCAGAATCCAACACG
MEF2D F	CAAGCTGTTCCAGTATGCCAG
MEF2D R	AAGGGATGATGTCACCAGGG
HDAC4 F	GAGAGAATTCTGCTAGCAATGAGCTCCCAA
HDAC4 R	GAGACTCGAGCTATGCAGGTTCCAAGGGCAGTGA
SRF F	GTCCCCATGCAGTGATGTATG
SRF R	GTAGCTCGGTGAGGTTGCTG

REFERENCES

1. **Ambros, V.** 2004. The functions of animal microRNAs. Nature **431:**350-355.

2. **Bartel, D. P.** 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell **116:**281-297.

3. **Blau, H. M., G. K. Pavlath, E. C. Hardeman, C. P. Chiu, L. Silberstein, S. G. Webster, S. C. Miller, and C. Webster.** 1985. Plasticity of the differentiated state. Science **230:**758-766.

4. **Brown, D. D., S. N. Martz, O. Binder, S. C. Goetz, B. M. Price, J. C. Smith, and F. L. Conlon.** 2005. Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development **132:**553-563.

5. **Cao, D., Z. Wang, C. L. Zhang, J. Oh, W. Xing, S. Li, J. A. Richardson, D. Z. Wang, and E. N. Olson.** 2005. Modulation of smooth muscle gene expression by association of histone acetyltransferases and deacetylases with myocardin. Mol Cell Biol **25:**364-376.

6. **Chen, C. Z., L. Li, H. F. Lodish, and D. P. Bartel.** 2004. MicroRNAs modulate hematopoietic lineage differentiation. Science **303:**83-86.

7. **Conlon, F. L., S. G. Sedgwick, K. M. Weston, and J. C. Smith.** 1996. Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. Development **122:**2427-2435.

8. **Giraldez, A. J., R. M. Cinalli, M. E. Glasner, A. J. Enright, J. M. Thomson, S. Baskerville, S. M. Hammond, D. P. Bartel, and A. F. Schier.** 2005. MicroRNAs regulate brain morphogenesis in zebrafish. Science **308:**833-838.

9. **He, L., J. M. Thomson, M. T. Hemann, E. Hernando-Monge, D. Mu, S. Goodson, S. Powers, C. Cordon-Cardo, S. W. Lowe, G. J. Hannon, and S. M. Hammond.** 2005. A microRNA polycistron as a potential human oncogene. Nature **435:**828-833.

10. **Hutvagner, G., M. J. Simard, C. C. Mello, and P. D. Zamore.** 2004. Sequence-specific inhibition of small RNA function. PLoS Biol **2:**E98.

11. **Kiriakidou, M., P. T. Nelson, A. Kouranov, P. Fitziev, C. Bouyioukos, Z. Mourelatos, and A. Hatzigeorgiou.** 2004. A combined computational-experimental approach predicts human microRNA targets. Genes Dev **18:**1165-1178.

12. **Krek, A., D. Grun, M. N. Poy, R. Wolf, L. Rosenberg, E. J. Epstein, P. MacMenamin, I. da Piedade, K. C. Gunsalus, M. Stoffel, and N. Rajewsky.** 2005. Combinatorial microRNA target predictions. Nat Genet **37:**495-500.

13. **Kroll, K. L., and E. Amaya.** 1996. Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. Development **122:**3173-3183.

14. **Lee, R. C., and V. Ambros.** 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. Science **294:**862-864.

15. **Lee, R. C., R. L. Feinbaum, and V. Ambros.** 1993. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell **75:**843-854.

16. **Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, and C. B. Burge.** 2003. Prediction of mammalian microRNA targets. Cell **115:**787-798.

17. **Li, S., M. P. Czubryt, J. McAnally, R. Bassel-Duby, J. A. Richardson, F. F. Wiebel, A. Nordheim, and E. N. Olson.** 2005. Requirement for serum response factor for skeletal muscle growth and maturation revealed by tissue-specific gene deletion in mice. Proc Natl Acad Sci U S A **102:**1082-1087.

18. **Lu, J., T. A. McKinsey, C. L. Zhang, and E. N. Olson.** 2000. Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. Mol Cell **6:**233-244.

19. **Mansfield, J. H., B. D. Harfe, R. Nissen, J. Obenauer, J. Srineel, A. Chaudhuri, R. Farzan-Kashani, M. Zuker, A. E. Pasquinelli, G. Ruvkun, P. A. Sharp, C. J. Tabin, and M. T. McManus.** 2004. MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat Genet **36:**1079-1083.

20. **McKinsey, T. A., C. L. Zhang, J. Lu, and E. N. Olson.** 2000. Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. Nature **408:**106-111.

21. **Meister, G., M. Landthaler, Y. Dorsett, and T. Tuschl.** 2004. Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA **10:**544-550.

22. **Soulez, M., C. G. Rouviere, P. Chafey, D. Hentzen, M. Vandromme, N. Lautredou, N. Lamb, A. Kahn, and D. Tuil.** 1996. Growth and differentiation of C2 myogenic cells are dependent on serum response factor. Mol Cell Biol **16:**6065-6074.

23. **Thomson, J. M., J. Parker, C. M. Perou, and S. M. Hammond.** 2004. A custom microarray platform for analysis of microRNA gene expression. Nat Methods **1:**47-53.

24. **Wang, D., P. S. Chang, Z. Wang, L. Sutherland, J. A. Richardson, E. Small, P. A. Krieg, and E. N. Olson.** 2001. Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. Cell **105:**851-862.

25. **Wang, D., R. Passier, Z. P. Liu, C. H. Shin, Z. Wang, S. Li, L. B. Sutherland, E. Small, P. A. Krieg, and E. N. Olson.** 2002. Regulation of cardiac growth and development by SRF and its cofactors. Cold Spring Harb Symp Quant Biol **67:**97-105.

26. **Wienholds, E., W. P. Kloosterman, E. Miska, E. Alvarez-Saavedra, E. Berezikov, E. de Bruijn, H. R. Horvitz, S. Kauppinen, and R. H. Plasterk.** 2005. MicroRNA expression in zebrafish embryonic development. Science **309:**310-311.

27. **Wightman, B., I. Ha, and G. Ruvkun.** 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell **75:**855-862.

28. **Zhao, Y., E. Samal, and D. Srivastava.** 2005. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. Nature **436:**214-220.

CHAPTER 6

GENERAL DISCUSSION

One of the most vital challenges in the field of cardiovascular development, and in developmental biology in general, is to define the precise mechanisms governing the restricted spatial and temporal expression of developmental genes. It is becoming clear, however, that complex regulatory networks, the combinatorial interactions of transcription factors, and *cis*-acting regulatory elements throughout the genome, are all key to the control of this regulation and have direct impacts on human disease. In order to fully address these questions, it will be imperative to take a two-pronged approach by which the regulatory elements of individual genes as well as larger signaling networks as a whole are examined in tandem, as well as to work to further develop the model systems available for such studies. The work described in this dissertation examines the connection between specific transcription factors and human disease, begins to characterize the transcriptional regulation of the cardiac transcription factor *Tbx20* and two microRNAs involved in muscle development, and demonstrates the importance and utility of the *X. tropicalis* system through the analysis of the conservation of four T-box genes, (8, 22, 33).

Cardiac transcription factors and human disease

In 1968, classical studies of congenital heart disease led to a hypothesis of multifactorial inheritance suggesting that the etiology of these cardiac defects was based on both genetic and environmental factors (26). Since that time, additional environmental factors have been identified, however, it has become increasingly clear that there is a significant genetic contribution to the occurrence of congenital heart disease (16). For this reason, current studies are aimed at carefully deciphering the signaling pathways that regulate early development and understanding how mutations in individual genes in these pathways lead to human pathologies such as congenital heart defects. As described in Chapter 2, mutations or deletions in a number of cardiac transcription factors have been associated with human congenital heart defects (22). Interestingly, of these factors, over one-third are members of the T-box family of transcription factors. For example, in 22q11 deletion syndrome, which represents the most common genetic deletion syndrome in humans, results in part from the deletion of *Tbx1*. DiGeorge syndrome is characterized by congenital heart defects such as tetralogy of Fallot, interruption of the aortic arch type B, ventricular septal defects, pulmonary atresia, or persistent truncus ateriosus (Figure 2.2)(42). In addition, over 30 mutations in the *Tbx5* locus have been linked to Holt-Oram syndrome, an autosomal dominant condition characterized by upper limb defects and cardiac defects ranging from septal defects to tetralogy of Fallot and hypoplastic left heart syndrome (Figure 2.2)(24). It was not until recently, however, that nonsense and missense mutations in the *Tbx20* locus were identified that result in a wide spectrum of developmental defects that includes septal, chamber and valve defects (19). It is interesting to note that, in a study of 352 individuals affected by congenital heart defects (CHDs), 0.6% of the patients carried a

mutation in the *Tbx20* gene, a rate similar to that of *Nkx2.5* mutations (11, 19, 23). More importantly, however, approximately 5% of all individuals with a family history of CHDs carried *Tbx20* mutations, suggesting that *Tbx20* may prove to be a key target of genetic screening in families afflicted with CHDs (19).

The continuing characterization of mutations and deletions in cardiac transcription factors and their association with a wide range of CHDs is beginning to emphasize that disruptions within just a single gene may lead to a vast array of defects that may or may not be consistent between individual patients. On the other hand, our continued understanding of the interplay between factors in developmental regulatory networks suggests that disruptions in multiple genes may result in similar clinical presentation. Both the diverse clinical presentations of patients with *Tbx20* mutations and the complex phenotypes resulting from TBX20 knock down in mouse and frog support these ideas (6, 35, 39, 40). For this reason, it will be important to identify any additional mutations within the *Tbx20* locus that could be at the root of other clinical manifestations. Human *Tbx20* maps to chromosome 7p14, and a search for pathologies with similar phenotypes to *Tbx20* knockdown, and that map to chromosome 7p14, reveals at least two potential candidates. First, a recent study by Bhuiyan *et al*. maps an early onset lethal form of catecholaminergic polymorphic ventricular tachycardia (CPVT) to chromosome 7p14-22 (2). CPVT is characterized by a reproducible form of ventricular tachycardia, induced by physical activity or stress, and is one of the most prevalent causes of sudden cardiac death in children (10, 21). While initial analyses of the *Tbx20* locus reveal no significant sequence variations, only coding exons and exon-intron boundaries were examined (2). In a similar manner, six other cardiac genes that map to this region were also found to contain no sequence alterations. It is possible that mutations in

regulatory regions of one or multiple gene(s) are responsible for the manifestation of this phenotype, leaving *Tbx20* as a potential candidate. Second, studies have demonstrated a requirement for TBX20 not only in cardiac development, but also in proper development of the nervous system, specifically of the motor neurons (37, 40). Loss of TBX20 leads to a complete loss of motor function in *Xenopus*, and to the loss of facial, trigeminal and vestibuloacoustic neuronal migration in mice (5, 37). For this reason, an investigation of hereditary neuropathies that map to the *Tbx20* locus is also warranted. One example is the inherited neuropathy axonal Charcot-Marie-Tooth disease (CMT2D), which maps to a region of chromosome 7p14 (12). Charcot-Marie-Tooth disease is one of the most common genetic diseases in humans and is characterized by a group of hereditary motor and sensory neuropathies (36). For these reasons, it possible that mutations in *Tbx20* could play a role in the manifestation of this disease, in addition to the congenital heart defects with which they have already been associated.

Cis-regulatory modules, *Tbx20* and the heart

As suggested above, many stages of organogenesis are controlled by a variety of interacting transcriptional regulatory networks. These networks involve the precise expression of genes and the proteins for which they encode in a spatially and temporally restricted pattern. The mechanism by which these transcriptional events are governed involves the interaction of transcription factors with target DNA sites that act as enhancers, silencers or insulators. These DNA sites, referred to as *cis*-regulatory elements, typically cover a few hundred base pairs and contain binding sites for multiple transcription factors (1, 14). Interestingly, comparative analysis of the sequences of *cis*-regulatory systems from

multiple species and diverse phyla reveals that there are a number of general characteristics shared by all of these systems (1). Arguably, the most striking characteristic of these regulatory regions is their high complexity. Generally, in a given *cis*-regulatory region, spanning less than 500bp, binding sites for positive and negative regulators, as well as their co-factors are found. A basic comparison of individual modules for a diverse array of developmental genes suggests that the complexity of an individual element responsible for the tissue-specific expression of a gene is approximately 4-8 different interactions with regulatory factors, and an average of five (1). Furthermore, it appears that these factors are chemically diverse along any one regulatory element in that no element has been described that is bound by only one protein type, such as homeodomain or Zn finger proteins (1).

Chapter 4 describes the identification of a cardiac *cis*-regulatory element for the transcription factor *Tbx20*. Briefly, a 334bp element was characterized and shown to be sufficient to drive the expression of *Tbx20* throughout the heart at specific phase of cardiac development (Figure 4.3). This element contains multiple conserved binding site sequences, and is bound by SMAD1, indicating that it is activated by BMP signaling (Figure 4.5-4.7). While this is the first step towards understanding the unique expression pattern of *Tbx20* in the heart, and adds an additional link to the ever-growing cardiac transcriptional network, it is most likely only the beginning of our understanding of *Tbx20* regulation. Based on the comparative analyses of *cis*-regulatory elements across species, and on the focused characterization of the regulation of other cardiac transcription factors such as *Nkx2.5* or *cardiac α-actin*, it is likely that *Tbx20* expression involves additional factors and additional regulatory elements. There are a number of attractive candidates based on expression patterns in the embryo, regulation of co-expressed cardiac genes, and preliminary studies.

One such candidate is serum response factor (SRF), a transcription factor involved in the regulation of a variety of muscle and cardiac genes (30). Specifically, SRF, in combination with *Nkx2.5*, has been shown to be required for the proper regulation of cardiac genes such as the *cardiac actin* promoter (7, 25). Additionally, conditional ablation of SRF in α-myosin heavy chain expressing cells results in reduced expression of cardiac genes such as atrial naturetic factor and cardiac, skeletal and smooth muscle α-actin (25). Preliminary studies in our lab suggest that *Tbx20*, through the defined 334bp regulatory element, is also positively regulated by SRF signaling in a dose dependent manner (Figure 6.1). This data, in addition to the expression of SRF in the heart, suggest that SRF is required for proper *Tbx20* expression in combination with BMP signaling. It is also possible that another cardiac transcription factor, GATA4, may be involved in *Tbx20* regulation. GATA4 is a zinc finger transcription factor that has been shown to bind sequence specific DNA elements in the regulation of cardiac genes such as *Nkx2.5* (3, 32). In a similar manner to *Tbx20*, GATA4 is required for proper cardiac development and is expressed from the early stages of cardiac development in the two bilateral patches of migrating cardiac primordia, through adulthood, in the fully developed adult heart (17, 20, 29). Sequence analysis using algorithms such as ConSite, TRANSFAC and JASPAR has identified multiple conserved GATA binding sites within the 334bp regulatory element of *Tbx20*. While GATA4 fails to significantly induce *Tbx20* expression in preliminary *in vitro* transcriptional assays, it is possible that it is required in combination with other transcription factors for the proper regulation of *Tbx20* (Figure 4.5). For example, GATA4 has been shown to act as a mutual co-factor in the activation of *Nkx2.5* in the heart through association with Smad1/4 on an upstream enhancer element (3). From this, it will be interesting to examine the potential interactions of GATA4

with other proteins in the activation of *Tbx20* in the heart. Most likely, it is the combinatorial activity of major signaling cascades such as BMP and SRF, in addition to other regulatory factors such as GATA4 at precise times and locations during development, that results in the unique expression pattern of *Tbx20* in the developing embryo.

X. tropicalis as a model for studies of development and disease

Much of our current understanding of gene regulation described both here and by others is a direct result of work in model systems such as *Xenopus*. The utility of the *Xenopus* model in studies of early development, including cardiogenesis, has been appreciated and expanding since the 1950s. Specifically, *X. laevis* has been well established and it is widely utilized in classical and molecular embryological studies. However, a number of drawbacks, including its tetraploid genome and long generation time, have prevented its use in genetic analyses. In light of this, the closely related *X. tropicalis* system is now emerging as a useful tool for developmental genetics. Unlike its counterpart, *X. tropicalis* is a truly diploid species with an available annotated genome sequence (Joint Genome Institute), and a short generation time, lending credence to its use in genetic studies (34). In Chapter 2, I describe an identification and structural analysis of four *X. tropicalis* cDNAs corresponding to *Tbx1*, *Tbx2*, *Tbx5* and *Tbx20*, a comparison of their identity and similarity across species, and a complete characterization of their expression patterns in *X. tropicalis*. Notably, these studies demonstrate the exceptional conservation of the sequence and expression of four key transcription factors between *X. tropicalis* and other vertebrate species. For example, the *Tbx20* sequence is greater than 97% identical between *X. tropicalis* and the well-characterized *X. laevis* model, and the RNA is expressed in nearly

identical temporal and spatial patterns in the cement gland, heart, eye and nervous system (33). From this, it is evident that *X. tropicalis* can be utilized in concert with the established *X. laevis* system, and that we should therefore focus our efforts in a few key directions in an effort to make it fully applicable to studies of human disease. First, it will be important to continue to work towards the optimization of forward and reverse genetic screens in *X. tropicalis*. Forward screens have utilized multiple approaches to disrupt gene function including N-ethyl N-nitrosourea treatment of sperm or morpholino oligonucleotide libraries, and are moving towards the ability to identify novel genes that regulate development (13, 31). In a similar manner, a number of reverse genetic screens have utilized methods such as TILLING (Targeting Induced Local Lesions in Genomes) to detect mutations in specific genes (38). While each of these methods has proven effective in small-scale scenarios, work is currently underway to achieve large-scale screening capabilities, which when reached, will reveal the true power of this system will be revealed. The second focus for research groups utilizing *X. tropicalis* will be the development of *Xenopus* models of human disease, as currently there are none available. While excellent mouse models of human disease do exist, *X. tropicalis* offers a number of distinct advantages including the accessibility of embryos due to external fertilization and development, and the ease and variety of available embryological techniques. In the past ten years, many additional tools have become available for the disruption and analysis of gene function in the frog. Utilizing techniques such as transgenesis, chemical mutagenesis and morpholino oligonucleotide injections that are widely utilized in *Xenopus*, efforts can be focused on the disruption of genes as they are identified in human disease (15, 27). With these models, the vast potential of screens directly

applicable to human disease, such as large-scale drug screening, can then be fully reached, as current efforts have shown great promise for widespread utility (41).

Future directions

Through the studies described in this dissertation and many others, it is becoming evident that the T-box transcription factor *Tbx20* is an essential link in an extensive signaling network that controls early heart development. This, in addition to the unique pattern of expression and an essential role in cardiac development, demonstrates the importance of gaining a more complete understanding of *Tbx20* regulation and function in development. To this end, we currently have a number of short- and long-term goals aimed at expanding this knowledge.

Further characterization of Tbx20 regulation

A major goal of this project in the immediate future is to more fully characterize *Tbx20* regulation in the developing embryo. The work described here characterizes a cardiac regulatory element that is acted upon by members of the BMP and SRF signaling pathways (Mandel *et al*., In Review). However, this element is not sufficient for the complete expression of *Tbx20*, in that it fails to drive expression of endogenous *Tbx20* in the early cardiac progenitor populations and nervous system (4, 33). To address this, we have begun to combine a comparative genomic approach with the available transgenic techniques in *Xenopus*. An initial comparison of the *X. tropicalis Tbx20* genomic locus with that of the human, mouse, chick, zebrafish and Fugu loci suggests that sequence conservation is a feasible approach to isolate putative regulatory regions *in silico*. From this, we have utilized

a transgenic system to identify 3 highly conserved intronic and 3'-UTR elements with ability to drive tissue-specific expression in the eye, brain, neural tube and rhombomeres (Figure 6.2, Data not shown). It will be important in the immediate future, however, to better characterize these putative elements and to expand this sequence analysis to extend both up- and downstream from the *Tbx20* coding region, in an effort to identify additional factors that may act at a greater distance from the *Tbx20* locus. In parallel with this genomic analysis, we will also take advantage of the transgenic technologies available in *Xenopus* to identify distal regulatory elements or elements that may work in a combinatorial manner to drive *Tbx20* expression. Reporter gene expression in transgenic *X. laevis* can be driven by recombineered bacterial artificial chromosomes (BACs) in an effort to isolate transcriptional regulatory elements over regions as large at 300Kb (18). A variety of *Xenopus* BAC libraries are currently available (C.H.O.R.I.), and we have identified a number of BAC clones that contain regions of the *Tbx20* coding region (Mandel *et al*., In Review). Utilizing the available transgenic techniques in *Xenopus*, it will be possible to quickly identify BACs that contain regions sufficient to drive specific *Tbx20* expression, and in combination with sequence comparisons across species, to efficiently define the precise regulatory elements across the *Tbx20* locus.

In addition to identifying regions within the *Tbx20* locus that are involved in its regulation, another goal of this project is to identify signaling pathways that are responsible for the proper developmental expression of *Tbx20*. Thus far, we have demonstrated a role for BMP signaling and SRF signaling in *Tbx20* regulation (Mandel *et al*., In Review). However, work to characterize the developmental regulation of other cardiac transcription factors, such as *Nkx2.5* and *Mef2c*, suggests that multiple signaling pathways are required to properly

regulate the spatially and temporally specific expression of individual cardiac genes (3, 9). To address this question for *Tbx20*, it will be possible to utilize a yeast one-hybrid approach to screen a *Xenopus* cardiac-specific cDNA library with the minimal *Tbx20* cardiac regulatory element. For this, we have created both cardiac-specific and embryo minus heart-specific *X. tropicalis* cDNA libraries from both stage 17 (tailbud) and stage 46 (tadpole) embryos (E.M.M., unpublished data). We will now utilize a modified linker scanning mutagenesis approach to identify the minimal regulatory element of *Tbx20* that is necessary for its expression in the heart. This element will then be used to screen the cardiac specific cDNA libraries to identify any other proteins that directly interact with the *Tbx20* regulatory element, thus giving us insight into any additional signaling pathways that may be involved in proper *Tbx20* expression. Additionally, the availability of this technique and the establishment of this system for use with *Xenopus* cDNA libraries and *Tbx20* will allow for the identification of signaling factors that interact with additional regulatory elements as they are identified, thus expanding our understanding of regulated gene expression during development.

Identification of downstream TBX20 targets

While the immediate goals of this project focus on the identification of upstream regulatory factors of *Tbx20*, it will also be necessary, in the longer term, to identify the downstream targets of TBX20 in order to piece together a complete transcriptional regulatory network. To date, TBX20 has been suggested to act upstream of a small number of cardiac and neural factors, including *Tbx2*, *Isl2*, and *Hb9* (35, 39, 40). However, given its requirement for proper heart and neural development, we hypothesize that TBX20 is

involved in the proper regulation of a wide panel of cardiac and neural factors. To identify additional downstream targets of TBX20, we will proceed with two parallel approaches, both of which take advantage of morpholino oligonucleotide knock down of TBX20, a technique which has been previously described in *Xenopus* (5). First, we will take a candidate-based approach by which we will examine differences in the expression patterns of candidate genes between wildtype embryos and those depleted of TBX20 expression. Preliminary work in the neural tube demonstrates that TBX20 loss leads to an expansion of the motor neuron marker, MNR2, and a reduction of the interneuron marker, EN1, suggesting that TBX20 may lie upstream of these factors (Figure 6.3; E.M.M., unpublished data). In parallel, we will utilize a microarray approach to identify genes that are differentially regulated upon knock down of TBX20. RNA will be isolated from wildtype or TBX20-depleted tissue (heart or neural tube) to screen the currently available *Xenopus* microarray chips (Affymetrix). Similar methods have proven to be efficient and effective for the identification of target genes in *Xenopus*, and will likely prove effective in our efforts to place TBX20 within a large transcriptional regulatory network (28).

Closing remarks

The work described within this dissertation represents a significant advance in our understanding of the signaling mechanisms and early events of cardiac development. Specifically, this work focuses on the T-box transcription factor *Tbx20*, and its expression, function and regulation in the developing heart. By defining the mechanisms of transcriptional regulation of individual cardiac factors that have been shown to be required at early stages of development, we can further our understanding of the molecular program that

regulates early organogenesis. Additionally, a further understanding of specific regulatory systems will not only provide insight in to the inner workings of individual signaling events, but will also indicated the specific interconnections between signaling cascades. Such studies are imperative to the field of cardiovascular biology in that alterations of early signaling events have been shown to underlie the manifestation of many human congenital diseases, and pose as a significant threat of human mortality. It is our hope that the work described herein provides useful insight into the unknown signaling events involved in early cardiogenesis and may eventually advance our understanding of human disease.

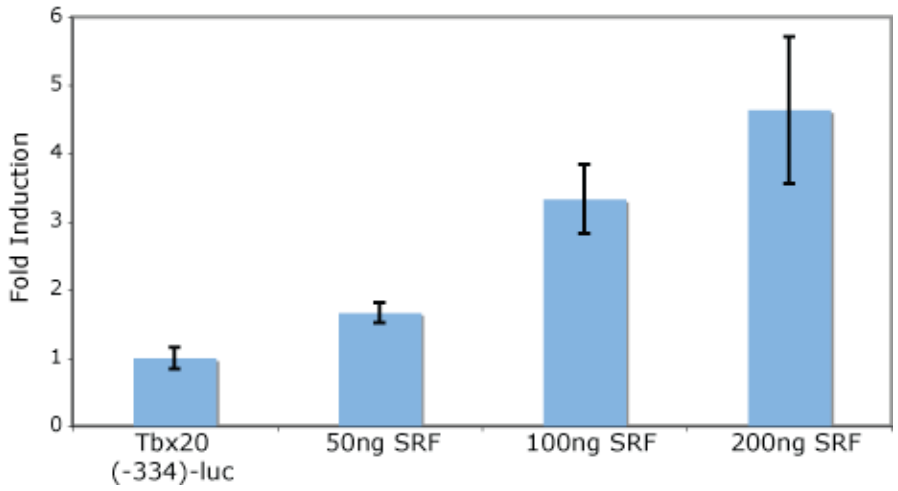


Figure 6.1. A 334bp 5' regulatory element of *Tbx20* is activated by SRF signaling. A luciferase reporter controlled by a 334bp 5' regulatory element is activated by increasing doses of SRF when co-transfected into COS7 cells as previously described (see chapter 4). Fold induction reflects changes in induction relative to induction of the reporter alone. Error bars represent standard error of 3 replicates.

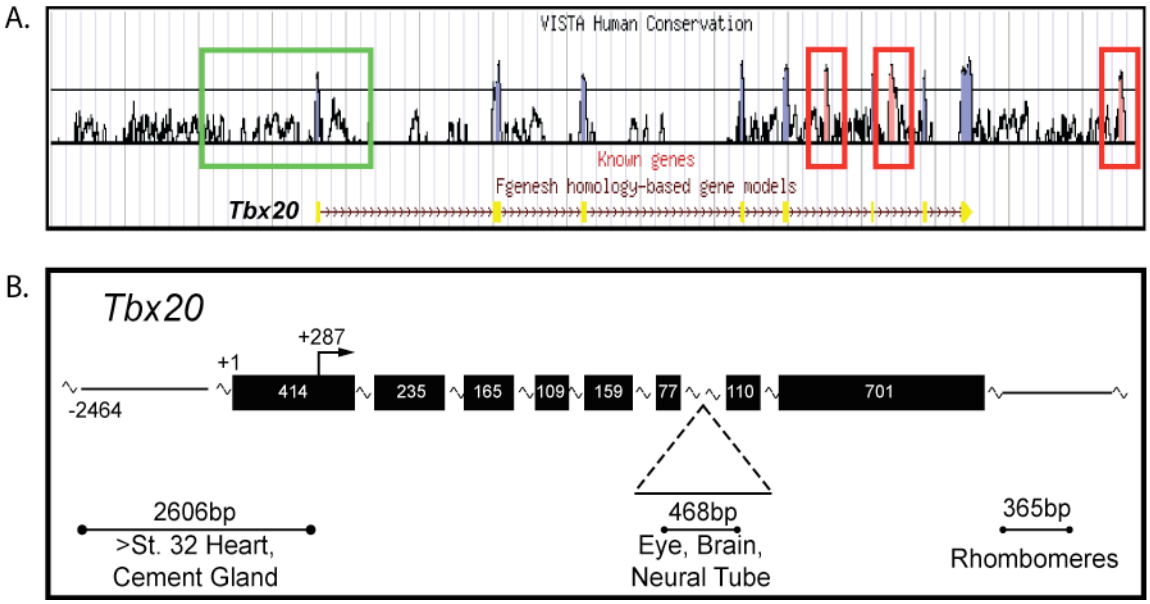


Figure 6.2. Evolutionarily conserved sequences drive tissue-specific expression of *Tbx20*. A, VISTA alignment and conservation between *X. tropicalis* and *H. sapien Tbx20*. Peaks represent the level of conservation between species using a 100bp window. Blue peaks denote coding sequence, while red peaks represent conservation of non-coding regions. The upper line represents a threshold of 63% conservation. The 2464bp cardiac regulatory element described in chapter 4 is boxed in green, while 3 putative regulatory elements are boxed in red. B, Schematic representation of the *Tbx20* locus and the location of 3 tissue-specific regulatory elements as determined by a transgenic analysis in *X. laevis*. The eight exons are represented by solid black boxes with the base pair length noted. Wavy lines represent non-coding sequences.

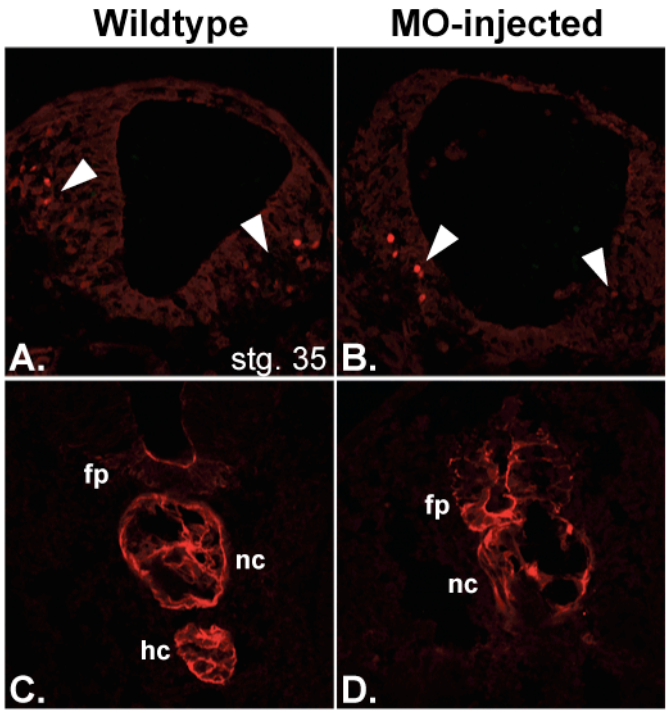


Figure 6.3. TBX20 knock down leads to changes in EN1 and MNR2 expression. Morpholino oligonucleotides (MO) against *Tbx20* were injected into single cell embryos as previously described to block translation of TBX20 (5). Wildtype and MO-injected embryos were fixed and sectioned for immunohistochemistry at stage 35 (see chapter 4) using α-EN-1 and α-Mnr2 primary antibodies (Developmental Studies Hybridoma Bank) and an α-mouse-cy3 secondary antibody. A-B, EN-1 expression marks a subset of interneurons at the lateral edge of the neural tube (arrowheads) in wildtype (A) and MO-injected (B) embryos. C-D, Mnr2 is expressed in the floor plate (fp), notochord (nc) and hypochord (hc) of wildtype (C) and MO-injected (D) embryos.

REFERENCES

1. **Arnone, M. I., and E. H. Davidson.** 1997. The hardwiring of development: organization and function of genomic regulatory systems. Development **124:**1851-1864.

2. **Bhuiyan, Z. A., M. A. Hamdan, E. T. Shamsi, A. V. Postma, M. M. Mannens, A. A. Wilde, and L. Al-Gazali.** 2007. A novel early onset lethal form of catecholaminergic polymorphic ventricular tachycardia maps to chromosome 7p14-p22. J Cardiovasc Electrophysiol **18:**1060-1066.

3. **Brown, C. O., 3rd, X. Chi, E. Garcia-Gras, M. Shirai, X. H. Feng, and R. J. Schwartz.** 2004. The cardiac determination factor, *Nkx2-5*, is activated by mutual cofactors GATA-4 and Smad1/4 via a novel upstream enhancer. J Biol Chem **279:**10659-10669.

4. **Brown, D. D., O. Binder, M. Pagratis, B. A. Parr, and F. L. Conlon.** 2003. Developmental expression of the *Xenopus laevis Tbx20* orthologue. Dev Genes Evol **212:**604-607.

5. **Brown, D. D., S. N. Martz, O. Binder, S. C. Goetz, B. M. J. Price, J. C. Smith, and F. L. Conlon.** 2005. Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development **132:**553-563.

6. **Cai, C. L., W. Zhou, L. Yang, L. Bu, Y. Qyang, X. Zhang, X. Li, M. G. Rosenfeld, J. Chen, and S. Evans.** 2005. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. Development. **132**:2475-2487.

7. **Chen, C. Y., J. Croissant, M. Majesky, S. Topouzis, T. McQuinn, M. J. Frankovsky, and R. J. Schwartz.** 1996. Activation of the *cardiac alpha-actin* promoter depends upon serum response factor, Tinman homologue, Nkx-2.5, and intact serum response elements. Dev Genet **19:**119-130.

8. **Chen, J. F., E. M. Mandel, J. M. Thomson, Q. Wu, T. E. Callis, S. M. Hammond, F. L. Conlon, and D. Z. Wang.** 2006. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet **38:**228-233.

9. **Dodou, E., M. P. Verzi, J. P. Anderson, S. M. Xu, and B. L. Black.** 2004. *Mef2c* is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. Development **131:**3931-3942.

10. **Eisenberg, S. J., M. M. Scheinman, N. K. Dullet, W. E. Finkbeiner, J. C. Griffin, M. Eldar, M. R. Franz, R. Gonzalez, A. H. Kadish, and M. D. Lesh.** 1995. Sudden cardiac death and polymorphous ventricular tachycardia in patients with normal QT intervals and normal systolic cardiac function. Am J Cardiol **75:**687-692.

11. **Elliott, D. A., E. P. Kirk, T. Yeoh, S. Chandar, F. McKenzie, P. Taylor, P. Grossfeld, D. Fatkin, O. Jones, P. Hayes, M. Feneley, and R. P. Harvey.** 2003. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. J Am Coll Cardiol **41:**2072-2076.

12. **Ellsworth, R. E., V. Ionasescu, C. Searby, V. C. Sheffield, V. V. Braden, T. A. Kucaba, J. D. McPherson, M. A. Marra, and E. D. Green.** 1999. The CMT2D locus: refined genetic position and construction of a bacterial clone-based physical map. Genome Res **9:**568-574.

13. **Goda, T., A. Abu-Daya, S. Carruthers, M. D. Clark, D. L. Stemple, and L. B. Zimmerman.** 2006. Genetic screens for mutations affecting development of *Xenopus tropicalis*. PLoS Genet **2:**e91.

14. **Habets, P. E., A. F. Moorman, and V. M. Christoffels.** 2003. Regulatory modules in the developing heart. Cardiovasc Res **58:**246-263.

15. **Heasman, J., M. Kofron, and C. Wylie.** 2000. Beta-catenin signaling activity dissected in the early *Xenopus* embryo: a novel antisense approach. Dev Biol **222:**124-134.

16. **Jenkins, K. J., A. Correa, J. A. Feinstein, L. Botto, A. E. Britt, S. R. Daniels, M. Elixson, C. A. Warnes, and C. L. Webb.** 2007. Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation **115:**2995-3014.

17. **Jiang, Y., and T. Evans.** 1996. The Xenopus GATA-4/5/6 genes are associated with cardiac specification and can regulate cardiac-specific transcription during embryogenesis. Dev Biol **174:**258-270.

18. **Kelly, L. E., B. E. Davy, N. F. Berbari, M. L. Robinson, and H. M. El-Hodiri.** 2005. Recombineered *Xenopus tropicalis* BAC expresses a GFP reporter under the control of *Arx* transcriptional regulatory elements in transgenic *Xenopus laevis* embryos. Genesis **41:**185-191.

19. **Kirk, E. P., M. Sunde, M. W. Costa, S. A. Rankin, O. Wolstein, M. L. Castro, T. L. Butler, C. Hyun, G. Guo, R. Otway, J. P. Mackay, L. B. Waddell, A. D. Cole, C. Hayward, A. Keogh, P. Macdonald, L. Griffiths, D. Fatkin, G. F. Sholler, A. M. Zorn, M. P. Feneley, D. S. Winlaw, and R. P. Harvey.** 2007. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. Am J Hum Genet **81:**280-291.

20. **Laverriere, A. C., C. MacNeill, C. Mueller, R. E. Poelmann, J. B. Burch, and T. Evans.** 1994. GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. J Biol Chem **269:**23177-23184.

21. **Leenhardt, A., V. Lucet, I. Denjoy, F. Grau, D. D. Ngoc, and P. Coumel.** 1995. Catecholaminergic polymorphic ventricular tachycardia in children. A 7-year follow-up of 21 patients. Circulation **91:**1512-1519.

22. **Mandel, E. M., T.E. Callis, D.Z. Wang, F.L. Conlon.** 2005. Transcriptional Mechanisms of Congenital Heart Disease. Drug Discovery Today: Disease Mechanisms **2:**33-38.

23. **McElhinney, D. B., E. Geiger, J. Blinder, D. W. Benson, and E. Goldmuntz.** 2003. NKX2.5 mutations in patients with congenital heart disease. J Am Coll Cardiol **42:**1650-1655.

24. **Mori, A. D., and B. G. Bruneau.** 2004. TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol **19:**211-215.

25. **Niu, Z., W. Yu, S. X. Zhang, M. Barron, N. S. Belaguli, M. D. Schneider, M. Parmacek, A. Nordheim, and R. J. Schwartz.** 2005. Conditional mutagenesis of the murine *serum response factor* gene blocks cardiogenesis and the transcription of downstream gene targets. J Biol Chem **280:**32531-32538.

26. **Nora, J. J.** 1968. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. Circulation **38:**604-617.

27. **Ogino, H., W. B. McConnell, and R. M. Grainger.** 2006. Highly efficient transgenesis in *Xenopus tropicalis* using I-SceI meganuclease. Mech Dev **123:**103-113.

28. **Park, E. C., T. Hayata, K. W. Cho, and J. K. Han.** 2007. *Xenopus* cDNA microarray identification of genes with endodermal organ expression. Dev Dyn **236:**1633-1649.

29. **Peterkin, T., A. Gibson, M. Loose, and R. Patient.** 2005. The roles of GATA-4, -5 and -6 in vertebrate heart development. Semin Cell Dev Biol **16:**83-94.

30. **Posern, G., and R. Treisman.** 2006. Actin' together: serum response factor, its cofactors and the link to signal transduction. Trends Cell Biol **16:**588-596.

31. **Rana, A. A., C. Collart, M. J. Gilchrist, and J. C. Smith.** 2006. Defining synphenotype groups in *Xenopus tropicalis* by use of antisense morpholino oligonucleotides. PLoS Genet **2:**e193.

32. **Searcy, R. D., E. B. Vincent, C. M. Liberatore, and K. E. Yutzey.** 1998. A GATA-dependent *nkx-2.5* regulatory element activates early cardiac gene expression in transgenic mice. Development **125:**4461-4470.

33. **Showell, C., K. S. Christine, E. M. Mandel, and F. L. Conlon.** 2006. Developmental expression patterns of *Tbx1*, *Tbx2*, *Tbx5*, and *Tbx20* in *Xenopus tropicalis*. Dev Dyn **235:**1623-1630.

34. **Showell, C., and F. L. Conlon.** 2007. Decoding development in *Xenopus tropicalis*. Genesis **45:**418-426.

35. **Singh, M. K., V. M. Christoffels, J. M. Dias, M. O. Trowe, M. Petry, K. Schuster-Gossler, A. Burger, J. Ericson, and A. Kispert.** 2005. Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. Development **132:**2697-2707.

36. **Skre, H.** 1974. Genetic and clinical aspects of Charcot-Marie-Tooth's disease. Clin Genet **6:**98-118.

37. **Song, M. R., R. Shirasaki, C. L. Cai, E. C. Ruiz, S. M. Evans, S. K. Lee, and S. L. Pfaff.** 2006. T-Box transcription factor Tbx20 regulates a genetic program for cranial motor neuron cell body migration. Development **133:**4945-4955.

38. **Stemple, D. L.** 2004. TILLING--a high-throughput harvest for functional genomics. Nat Rev Genet **5:**145-150.

39. **Stennard, F. A., M. W. Costa, D. Lai, C. Biben, M. B. Furtado, M. J. Solloway, D. J. McCulley, C. Leimena, J. I. Preis, S. L. Dunwoodie, D. E. Elliott, O. W. Prall, B. L. Black, D. Fatkin, and R. P. Harvey.** 2005. Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. Development **132**:2451-2462.

40. **Takeuchi, J. K., M. Mileikovskaia, K. Koshiba-Takeuchi, A. B. Heidt, A. D. Mori, E. P. Arruda, M. Gertsenstein, R. Georges, L. Davidson, R. Mo, C. C. Hui, R. M. Henkelman, M. Nemer, B. L. Black, A. Nagy, and B. G. Bruneau.** 2005. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. Development **132**:2463-2474.

41. **Tomlinson, M. L., R. A. Field, and G. N. Wheeler.** 2005. *Xenopus* as a model organism in developmental chemical genetic screens. Mol Biosyst **1:**223-228.

42. **Yamagishi, H., and D. Srivastava.** 2003. Unraveling the genetic and developmental mysteries of 22q11 deletion syndrome. Trends Mol Med **9:**383-389.