THE UNIVERSITY OF CHICAGO

DEFINING GENE REGULATORY NETWORKS IN CARDIAC RHYTHM AND DISEASE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON DEVELOPMENT, REGENERATION, AND STEM CELL BIOLOGY

BY

SONJA LAZAREVIC

CHICAGO, ILLINOIS

DECEMBER 2022

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my adviser and mentor, Dr. Ivan Moskowitz. I have appreciated his mentorship and support during my tenure in his laboratory. He has always been encouraging of my research endeavors, and ensured that I had all the available resources and collaborations to move our projects in an exciting direction. Most importantly, I knew I could reach out to him if I needed anything and he would always be there to help. I would like to thank the members of my committee, Dr. Barbara Kee, Dr. Megan McNerney, and Dr. Marcelo Nobrega, for their guidance and advice throughout the years. I would also like to acknowledge our amazing collaborators, Dr. David Park, Dr. Alexandre Colas, Dr. Alex Ruthenburg, and Dr. Sebastian Pott. They were instrumental in the advancement of our research projects, and I am thankful for the opportunity to work with each of them. I am also grateful to have had Dr. Rodney Dale as a mentor during the completion of my Master's degree. I learned so much during my time in his lab, and he was always supportive and encouraging of my scientific aspirations.

I am very appreciative of the past and present members of the Moskowitz lab that I have had the opportunity to work with. The Moskowitz lab has always had very talented and dedicated individuals, and I am fortunate to have worked with them. I have to acknowledge Zhezhen Wang, a talented bioinformatician in the lab that I collaborated closely with. We met almost every day and made the best team. I am appreciative of all her efforts that were instrumental in driving the bioinformatic analysis forward. Dr. Linsin Smith, a talented graduate student in Moskowitz lab, also deserves special acknowledgement as she has been an immense source of support for me during graduate school. She was always there to listen and provide the best insight into any problem, whether it was related to our research or any aspect of life. I am grateful to have worked with her, but more importantly I am thankful to have made an amazing friend. I would also like to

thank Dr. Brigitte Laforest, Dr. Alexander Guzzetta, Dr. Megan Rowton, Dr. Josh Theisen, Dr. Rajiv Nadadur, Dr. Ozanna Burnicka-Turek, Jessica Jacobs-Li, and Carlos Perez-Cervantes. They are brilliant individuals that I am privileged to have worked with, and they are one of the many reasons that I enjoy coming to the lab every day. I was also fortunate to mentor very talented undergraduates and technicians in the lab including, Margaret Gadek, Kaitlyn Shen, and Janvi Kukreja. They are very talented individuals, and I consider myself lucky to be a small part of their scientific endeavors. I am very excited to see the amazing things they will accomplish in the future. I would also like to Dr. Kohta Ikegami for his support and guidance. Even when he was very busy, his door was always open and he took the time to answer any questions .

Finally, I would like to thank my friends and family for their unwavering support during graduate school. My parents immigrated to the United States with nothing and built a beautiful life for our family. Without a doubt their love, support, and hard-work is the reason I am here today. Thank you to my siblings and their families for always being a pillar of support and unconditional love. There is nothing I enjoy more than watching our children grow up together. I would like to thank my husband, for his unconditional support. We got married the summer before starting graduate school at the University of Chicago, so we have really been on this journey together. We overcame a lot of obstacles, and we always tackled them together as a team. I am thankful to always have you by my side. Lastly, I want to thank my twins, Stefan and Nikola, who are my greatest accomplishment in life. They are both strong, kind, considerate, fearless, generous, and hard-working. I am forever lucky to be their Mom, and I am so excited to see what their future holds. Volim vas puno.

ABSTRACT

Atrial fibrillation (AF) is the most common type of cardiac arrhythmia, characterized by an irregular heartbeat due to uncoordinated electrical activity in the atria. The *Tbx5* locus has been implicated by genome wide association studies (GWAS), and our laboratory has previously shown that the absence of *Tbx5* leads to spontaneous and sustained AF. We will utilize this *Tbx5*-deficient mouse model to further characterize the regulatory mechanisms important for cardiac rhythm and the molecular mechanisms driving the pathophysiology of atrial fibrillation.

Cardiac rhythm is a dose-sensitive physiologic process, therefore stable gene expression is imperative for proper cardiac function. Precise gene regulation is reliant on feedback mechanisms to maintain accurate gene expression, and miRNAs are a canonical feedback mechanism for steady gene expression. In Chapter 1, our laboratory performed small RNA profiling of *Tbx5*-deficient mice to define candidate miRNAs involved in the TBX5-dependent gene regulatory network important for the maintenance of cardiac rhythm. Many of the miRNAs identified were interrogated in a high-throughput screen, which linked arrhythmogenic phenotypes to several miRNAs. Whole mouse and cellular electrophysiology studies were focused on a single miRNA candidate, miR-10b, and demonstrated its mis-regulation leads to atrial fibrillation susceptibility. This work supports a model where TBX5 regulates the expression of miRNAs critical for maintaining appropriate gene expression levels important in cardiac rhythm.

Atrial fibrillation also has a strong epidemiologic link with heart failure, as these two cardiac diseases are associated with increased incidence of each other. In Chapter 2, we revealed remarkable correlation between the differentially expressed genes in the atria of AF and heart failure mouse models, and identified shared transcription factors as candidates important in driving the pathophysiology of cardiac disease. We also identified the conservation of differential ncRNA

transcripts in both of these disparate disease models. Based on the knowledge that noncoding RNAs (ncRNAs) are transcribed from regulatory elements, it supports the paradigm of a common disease-specific gene regulatory network that mediates the physiologic consequences of disease. These differential non-coding RNA transcripts identified a TBX5-dependent candidate regulatory element downstream of *Klf15*, an important regulator in cardiac hypertrophy. We also identified a ncRNA transcript upstream of *Sox9* that is up-regulated in both disease models, and may uncover a disease-response essential for coping with atrial dysfunction. In summary, our studies have identified crucial transcriptional changes in atrial fibrillation and heart failure, along with the shared regulatory mechanisms critical in driving these changes.

Finally, in Chapter 4 we performed a time course deletion of *Tbx5* in an effort to identify the earliest transcriptional changes in this atrial fibrillation mouse model. Interestingly, the removal of *Tbx5* leads to an early response at Day 3 and Day 6, which is very different from the response at Day 10 and Day 17 of the time course. In an effort to identify significant gene expression changes between *Tbx5* KO and WT throughout the time course we performed maSigPro analysis, which has provided us with candidate genes important in driving the pathophysiology of atrial fibrillation.

CHAPTER 1: INTRODUCTION

The heart is an extraordinary organ that is composed of an assortment of cells that function synchronously to pump blood, which provides oxygen and nutrients to the organs and tissues within the body. A healthy heart is capable of pumping a sufficient amount of blood to meet the metabolic needs of the body. Deoxygenated blood enters the heart in the right atria, travels to the right ventricle, and then moves through the pulmonary arteries to the lungs. The lungs are crucial for re-oxygenating the blood, where it then flows through the pulmonary vein to the left atria. The blood subsequently travels from the left atria to the left ventricle, where it gets pumped through the aorta to provide oxygen and nutrients to the body.

Cardiac conduction

The coordinated movement of blood through the chambers of the heart is controlled by the cardiac conduction system and its organized electrical activity. An autonomous excitation signal originates from the sinoatrial (SA) node, propagates through the atria with cardiomyocyte depolarizations to contract and force blood to the ventricles. The electrical signal briefly pauses in the atrioventricular (AV) node to allow a sufficient amount of time for the ventricles to fill. The signal then travels down the bundle of His and Purkinje fibers for ventricular depolarization. The electrical signal in the heart propagates through the chambers due to the action potential of the cardiomyocytes. The resting potential of cardiomyocytes is ~-90mV, until depolarization leads to a positive voltage due to the influx of ions through sodium channels [1]–[3]. Calcium ions then enter the cell through L-type calcium channels, which results in the release of stored calcium ions from the ryanodine receptor, RYR2, on the sarcoplasmic reticulum [1]–[5]. The increased levels of intracellular calcium are instrumental for the contraction of the cardiomyocytes. The intracellular calcium is subsequently sequestered in the sarcoplasmic reticulum with the support

of SERCA2, sarco-endoplasmic reticulum Ca^{2+}-ATPase [1]–[5]. The final step of cardiac conduction is rapid repolarization with the closing of the L-type calcium channels and outflow of potassium ions [1]–[5]. The activity of these ion channels is crucial to maintain a coordinated electrical signal through the heart. This synchronized electrical activity is required to maintain a steady, rhythmic heartbeat in the heart; which is in turn required for appropriate systolic and diastolic contractions in the heart.

Atrial fibrillation

Unfortunately, there are a group of conditions that negatively affect the function of the heart required for proper blood flow, and these cardiovascular diseases are the leading cause of death in the United States and globally [6]. While cardiovascular diseases affect various aspects of the cardiovascular system, cardiac arrhythmias are a subset of cardiovascular diseases characterized by an abnormal heartbeat [7]. Atrial fibrillation (AF) is the most common type of cardiac arrhythmia, and it is characterized by abnormal, irregular heartbeats due to uncoordinated electrical activity in the atria. The irregular electrical activity in the atria can lead to the loss of synchronization with the ventricles. This is a particularly dangerous condition because it can result in pooling of blood in the atria, which can lead to blood clots and increased risk of strokes [8]. Individuals with atrial fibrillation are also more susceptible to comorbidities, including heart failure, dementia, and death [9], [10].

Atrial fibrillation is a growing epidemic that is estimated to affect 2.5-5.1 million people in the United States, but projections expect the incidence to more than double by 2050, which estimates around 5.6-12.1 million Americans will be affected by this condition [11], [12]. The progressively increasing rates of atrial fibrillation are not isolated to the United States, and global estimates of atrial fibrillation incidence are ~33.5 million individuals [13]. As the incidence of

atrial fibrillation continues to increase, the identification of novel therapeutic treatments is necessary, as many of the current treatments have significant side effects and can have short-term efficacy [9], [14]–[16]. Anticoagulation therapy, using blood thinners like warfarin, is commonly utilized in individuals with atrial fibrillation to prevent strokes, as these individuals are more susceptible to atrial blood clots [8], [9], [14]–[16]. Anti-arrhythmic drugs, like beta blockers or ion channel blockers, are also commonly used to control heart rate in order to maintain sinus rhythm. Finally, cardiac ablations are another common treatment for atrial fibrillation, and are accomplished by scarring areas of the heart to hinder abnormal electrical signals [9], [14]–[17]. The course of treatment for patients with atrial fibrillation is determined by the frequency of the AF episodes. Atrial fibrillation is a progressive condition, which is summarized by the common proverb 'AF begets AF'. Therefore, atrial fibrillation can be classified into categories based on the frequency of episodes. Paroxysmal atrial fibrillation is characterized by intermittent episodes that last less than seven days, and these episodes can eventually progress into persistent and permanent atrial fibrillation [9], [14], [16]. Electrical and structural remodeling occurs in atrial fibrillation, which can lead to progressively increasing frequency and duration of these symptoms. Anti-arrhythmic drugs are generally used as a primary approach to restore sinus rhythm, but when they are ineffective they are used in conjunction with cardiac ablations. While these treatments are able to mitigate the symptoms of the disease, many times they do not have long-term efficacy as patients revert back to AF. Therefore, an advanced understanding of atrial fibrillation pathophysiology will be critical for identifying more effective therapeutic treatments.

Heritability of atrial fibrillation pathophysiology

The pathophysiology of atrial fibrillation is believed to be composed of a trigger and an underlying substrate that propagates the abnormal signal in the atria [7], [18], [19]. In a healthy

heart the automaticity of the sinus node initiates the electrical signal that subsequently propagates through the heart. In atrial fibrillation there is a trigger caused by an ectopic depolarization in the atria [7], [18], [19]. Focal ectopic depolarizations can be caused by delayed afterdepolarizations (DADs) or early afterdepolarizations (EADs). DADs occur during the resting potential of the action potential and is generally caused by calcium handling abnormalities. EADs occur during the repolarization phase, and are generally caused by a prolonged repolarization of the action potential. While a trigger can be the initiating factor in atrial fibrillation, the perpetuation of the abnormal signal is due to the presence of a substrate from electrical or structural remodeling in the heart. For example, during cardiac disease structural remodeling processes include increased atrial fibrosis, which results in impaired cell communications and disrupted electrical activity in the heart. The presence of these two components is sufficient to propagate an abnormal electrical signal that leads to an irregular heartbeat.

Atrial fibrillation was originally thought to be a consequence of the structural remodeling that occurs in cardiac disease, but it is now clear there is genetic component to atrial fibrillation as well [20]–[25]. Atrial fibrillation can be inherited, therefore the importance of elucidating the risk loci of atrial fibrillation will be of the upmost importance as the incidence for atrial fibrillation continues to rise. Genome-wide association studies (GWAS) and familial studies have associated genetic variants to atrial fibrillation risk in cardiac transcription factors, ion channels, and cardiac structural genes [20]–[31]. A mutation in KCNQ1, a potassium channel, was one of the first risk loci identified in a familial AF linkage analysis [32]. A serine to glycine missense mutation resulted in a gain-of-function mutation that was present in family members with atrial fibrillation [32]. Subsequent mutations in the KCNQ1 gene have been identified to have an increased risk of atrial fibrillation [33]–[39]. Mutations have also been identified in other potassium channels, and sodium

channels as well, demonstrating the importance of ion channels in cardiac rhythm homeostasis [40]–[46]. GWAS have also been instrumental in identifying putative risk loci for atrial fibrillation. The first GWAS for atrial fibrillation identified SNPs on chromosome 4q25, which were located near PITX2, Paired Like Homeodomain 2 transcription factor [30]. Further studies have demonstrated that altered PITX2 expression results in atrial fibrillation susceptibility [47]–[49]. Subsequent GWAS analyses have identified many other risk loci for atrial fibrillation, and include transcription factors, gap junction proteins, and ion channels [26], [28], [30], [31], [50]–[57]. TBX5 (T-Box transcription factor 5), GATA4 (GATA binding protein 4), and NKX2.5 (NK2 homeobox 5) are examples of transcription factors that have risk loci for atrial fibrillation.

TBX5, GATA4, and NKX2.5 are cardiogenic transcription factors with important functions during heart development, and they are now known to have important roles in adult heart rhythm homeostasis. These transcription factors have been shown to physically interact to cooperatively regulate downstream gene expression [58]–[69]. GWAS studies have also implicated these transcription factors as risk loci for atrial fibrillation susceptibility [20], [23], [27]. Subsequent studies have demonstrated that GATA4 and NKX2.5 mutations are associated with familial atrial fibrillation [70]–[75]. A TBX5 variant (rs3825214) was identified in a GWAS for PR interval and QRS duration, clearly linking TBX5 with electrophysiologic parameters related to atrial fibrillation [76]. The association of this TBX5 SNP with atrial fibrillation was subsequently validated in a Chinese Han population [29]. Studies have since identified additional mutations in TBX5 that provided further evidence that TBX5 is involved in the pathogenesis of atrial fibrillation [77], [78]. TBX5 is also a transcription factor associated with Holt-Oram syndrome, an autosomal dominant condition characterized by upper limb deformities and cardiac defects [79]–[81]. The upper limb defects are variable, and can range from mild bone deformities in the hand to more

severe abnormalities like phocomelia [79]–[81]. The cardiac pathologic phenotypes in patients with Holt-Oram syndrome can vary, and include structural defects that result in congenital heart defects, including septal defects [79]–[81]. Holt-Oram syndrome can also result in conduction defects, including atrial fibrillation, atrioventricular block, and sinus node dysfunction [79]–[81]. Initial linkage analysis identified the gene responsible for Holt-Oram syndrome was located on chromosome 12, but the Brook and Seidman laboratories simultaneously published that mutations in TBX5 are the cause for this syndrome [79], [82], [83]. Quan Li Yu and colleagues identified insertion mutations that result in a reading frame shift and substitution mutations that result in a premature stop codon within the TBX5 gene [83]. Basson and colleagues identified two mutations in TBX5 for families with Holt-Oram syndrome, including a missense mutation for a premature stop codon [82]. A gain-of-function TBX5 mutation was subsequently identified in a family with atypical Holt-Oram syndrome and paroxysmal atrial fibrillation [84]. This family had mild skeletal defects, but only a few family members had congenital defects; instead, a majority of the family members developed paroxysmal atrial fibrillation at an early age [84]. This mutation resulted in enhanced DNA binding qualities that resulted in increased activation of TBX5 target genes, *Nppa* and *Cx40* [84]. These familial studies have been instrumental in identifying the cause of Holt-Oram syndrome, but also demonstrate the loss-of-function or gain-of-function mutations can lead to the syndrome, demonstrating the dose-dependent nature of TBX5 gene expression. While these GWAS and familial studies have been invaluable in identifying risk loci for atrial fibrillation, they also demonstrated the important of this transcription factor in maintaining cardiac rhythm homeostasis.

T-Box transcription factors

The T-box family of transcription factors are composed of proteins that bind DNA with a T-box DNA-binding motif, and they have a variety of critical roles in various aspects of development and disease [85]–[87]. The first T-box gene identified was Brachyury where it was identified to be a transcription factor important during mesoderm specification in development [88]–[90]. Subsequently, a total of seventeen T-box transcription factors were identified and categorized into five sub-families (T, Tbx1, Tbx2, Tbx6, and Tbr1) [85]–[87]. The members of the T-box transcription factor family bind to a AGGTGT consensus sequence [85], [86], [91]. The TBX2 sub-family is composed of the following genes: *Tbx2*, *Tbx3*, *Tbx4*, and *Tbx5*. This is an interesting sub-family, as phylogenetic analysis suggests they originated from a tandem duplication event [85], [86], [92], [93]. In the mouse genome, TBX5 and TBX3 are neighboring each other on chromosome 5, while TBX2 and TBX4 are located on chromosome 11. Several members of this sub-family have important roles in cardiac development, along with several family members of the TBX1 sub-family. A total of six T-box transcription factors are functionally important for the development of the heart, and include *Tbx2*, *Tbx3*, *Tbx5*, *Tbx1*, *Tbx18*, and *Tbx20*. *Tbx1* is expressed in the second heart field, involved in the development of the pharyngeal arch arteries, and participates in neural crest migration [86], [94]–[99]. The loss of *Tbx1* has been implicated in DiGeorge syndrome, and one of the characteristics defined by this syndrome is congenital heart defects [86], [94]–[99]. *Tbx2* and *Tbx3* are transcriptional repressors involved in the inhibiting the chamber myocardium during heart development [100]–[102]. *Tbx3* has also been shown to have a function role in the development of the cardiac conduction system [103]. *Tbx20* is critical for chamber specification through the repression of *Tbx2*, and mutations in the gene have been linked to congenital heart defects [104]–[107]. While it is clear that the T-box transcription

factors have a variety of functional roles in the developing heart, our work is focused on elucidating the functional role of *Tbx5*.

TBX5 is a transcription factor that is expressed in the developing heart, forelimb, and eye [108]–[110]. TBX5 expression is located in the forelimb during development, which is in contrast to TBX4 expression in the hindlimb [110]–[114]. It has since been shown that TBX5 is critical for the initiation of the forelimb bud formation through the direct activation of *Fgf10* in the early limb mesenchyme [114]. During cardiac development, TBX5 is highly expressed in the cardiac crescent, but it subsequently begins to favor the posterior portion of the heart tube [108]. As the heart undergoes looping, TBX5 is favored in the regions that will become the atria and left ventricle [108]. This regionalized expression of TBX5 will persist into adulthood, where there is a minimal expression in the right ventricle. TBX5 is important for the development of the cardiac conduction system, and *Tbx5* haploinsufficiency results in atrioventricular bundle and bundle branch defects [115]. TBX5 is also known to have a functional role in the adult cardiac conduction system, as it drives the expression of genes important for fast conduction in the ventricular conduction system, including *Scn5a* [116]. Interestingly, the balance of TBX5 and TBX3 expression in the cardiac conduction system determines whether the specialized electrophysiologic cells will be fast or slow conducting; a higher ratio of TBX5/TBX3 expression will result in fast conduction in the ventricular conduction system, while a lower ratio of TBX5/TBX3 expression will result in slow conduction indicative of the nodal cells [117]. TBX5 is a critical transcription factor for various aspects of heart development, and fortunately a mouse model was generated to explore the pathophysiology of *Tbx5*-deletion during development and in the adult heart.

Mouse models to evaluate Tbx5 function

A mouse model of *Tbx5*-deficiency has been developed to further elucidate the functional role of TBX5 in development and adult cardiac homeostasis. This *Tbx5*-deletion mouse model was created with the addition of loxP sites flanking the third exon of the TBX5 allele. In the presence of Cre recombinase, transcription of the floxed allele results in a shortened transcript that is inactive due to the deletion of the third exon. Homozygous deletion of TBX5 results in embryonic lethality, demonstrating the importance of the TBX5 during development. Interestingly, TBX5 haploinsufficiency mice with heterozygous *Tbx5* alleles causes limb and cardiac defects that are reminiscent of Holt-Oram syndrome in humans [64]. The limb defects in *Tbx5* heterozygous mice is characterized by elongated digits in the forelimb and hypoplastic bones in the wrist. *Tbx5* haploinsufficiency also results in enlarged hearts in the adult mouse, and other cardiac defects including atrial septal defects and conduction system abnormalities [64], [115].

Our laboratory has subsequently utilized the *Tbx5*-deficient mice to evaluate the effects of a conditional deletion of TBX5 in the adult mouse. The conditional deletion resulted in spontaneous, sustained atrial fibrillation (Figure 1.1) [118]. The *Tbx5*-deficient mice have all the hallmark characteristics of atrial fibrillation, including irregular heartbeat, prolonged action potential, and an absent p-wave [118]. The pacing of cardiomyocytes from *Tbx5*-deficient mice resulted in triggered action potentials, delayed afterdepolarizations (DADs), and early afterdepolarizations (EADs) due to calcium handling abnormalities. These irregularities are characteristic of triggers that initiate atrial fibrillation. In order to determine the genes regulated by TBX5, transcriptional profiling was performed on the left atria of control and *Tbx5*-deficient mice. The loss to *Tbx5* in the adult heart leads to decreased expression of genes important in cardiac rhythm, including calcium handling genes, potassium channels, sodium channels, and connexins. Interestingly, TBX5 also drives the expression of Pitx2, Paired Like Homeodomain 2, a repressive

transcription factor. TBX5 and PITX2 antagonistically regulated the expression of downstream calcium handling genes, including ryanodine (*Ryr2*) and SERCA (*Atp2a2*) [118]. *Ryr2* and *Atp2a2* are responsible for the flow of Ca^{2+} ions in and out of the sarcoplasmic reticulum (SR), therefore demonstrating the importance of TBX5 in calcium handling processes in the cardiomyocytes. A closer inspection of the electrophysiologic characteristics of these alleles in a heterozygous background reveals the importance of gene expression dosage for the expression of downstream cardiac rhythm genes. While the *Tbx5* heterozygotes ($Tbx5^{fl/+}$; $R26^{CreERT2}$) resulted in a prolonged action potential, instead the Pitx2 heterozygotes ($Pitx2^{fl/+}$; $R26^{CreERT2}$) caused shortened action potentials compared to the controls ($R26^{CreERT2}$). Both Tbx5 and Pitx2 haploinsufficiency led to atrial fibrillation susceptibility with pacing in the electrophysiologic studies. Interestingly, compound heterozygotes ($Tbx5^{fl/+}$; $Pitx2^{fl/+}$; $R26^{CreERT2}$) rescued the electrophysiologic abnormalities seen in the atrial fibrillation phenotype of the *Tbx5* heterozygote ($Tbx5^{fl/+}$; $R26^{CreERT2}$) [118]. The compound heterozygotes no longer displayed ectopic depolarizations or atrial fibrillation susceptibility, which demonstrates gene dosage is critical for cardiac rhythm genes, and critical for atrial rhythm stability. Overall, this study identified a novel atrial fibrillation mouse model, and identified an incoherent feed forward loop with TBX5 and PITX2 regulating the expression of dose-sensitive cardiac rhythm genes.

Our laboratory utilized this atrial fibrillation mouse model to further interrogate the regulatory mechanism important in controlling the electrical activity within the heart. TBX5 and GATA4 are two cardiac transcription factors that have been shown to cooperatively regulate downstream gene expression. Interestingly, TBX5 haploinsufficiency results in atrial fibrillation susceptibility due to calcium abnormalities from reduced SERCA function, but compound heterozygotes for TBX5 and GATA4 ($Tbx5^{fl/+}$; $Gata4^{fl/+}$; $R26^{CreERT2}$) rescue the arrhythmogenic

phenotype in the *Tbx5* heterozygotes (*Tbx5*$^{fl/+}$; *R26*^{CreERT2}) [119]. This study demonstrates that TBX5 and GATA4 are critical for maintaining homeostasis of the calcium signaling processes. In conjunction with our collaborators in the Weber lab, our laboratory continued to explore the effects of *Tbx5* on the electrical activity of the cardiomyocytes [120]. A closer inspection of the electrophysiology in the *Tbx5*-deficient mice revealed a decrease in SR load due to decreased SERCA function, and increased extrusion of Ca^{2+} with the sodium-calcium exchanger. Our laboratory has previously demonstrated that SERCA expression levels are regulated by TBX5 activation, but SERCA activity is also modulated by phospholamban (*Pln*), a negative regulator. Interestingly, the arrhythmogenic phenotype seen with *Tbx5* removal (*Tbx5*$^{fl/fl}$; *R26*$^{CreERT2)}$ is rescued in compound homozygotes with Pln removal (*Tbx5*$^{fl/fl}$; *Pln*$^{-/-}$; *R26*^{CreERT2}) [120]. The action potential duration and frequency of EADs/DADs were returned to baseline in the *Tbx5* and *Pln* double knockouts, and they were no longer susceptible to pacing-induced atrial fibrillation. These studies demonstrate that calcium handling irregularities are the cause of our atrial fibrillation mouse model. It also reveals the dose-sensitive nature of the calcium-handling genes for maintenance of cardiac rhythm homeostasis. While the dose-sensitive nature of gene expression is a well-established phenomenon, our laboratory has provided several examples of phenotypic rescue of the arrhythmogenic phenotype seen in our atrial fibrillation mouse model. We will continue to utilize this *Tbx5*-deletion mouse model to further characterize the pathophysiology of atrial fibrillation using a variety of genomic methods. Elucidating the transcriptional regulators and the regulatory mechanisms will illuminate the molecular mechanisms in maintaining cardiac rhythm homeostasis, and how it is perturbed in atrial fibrillation.

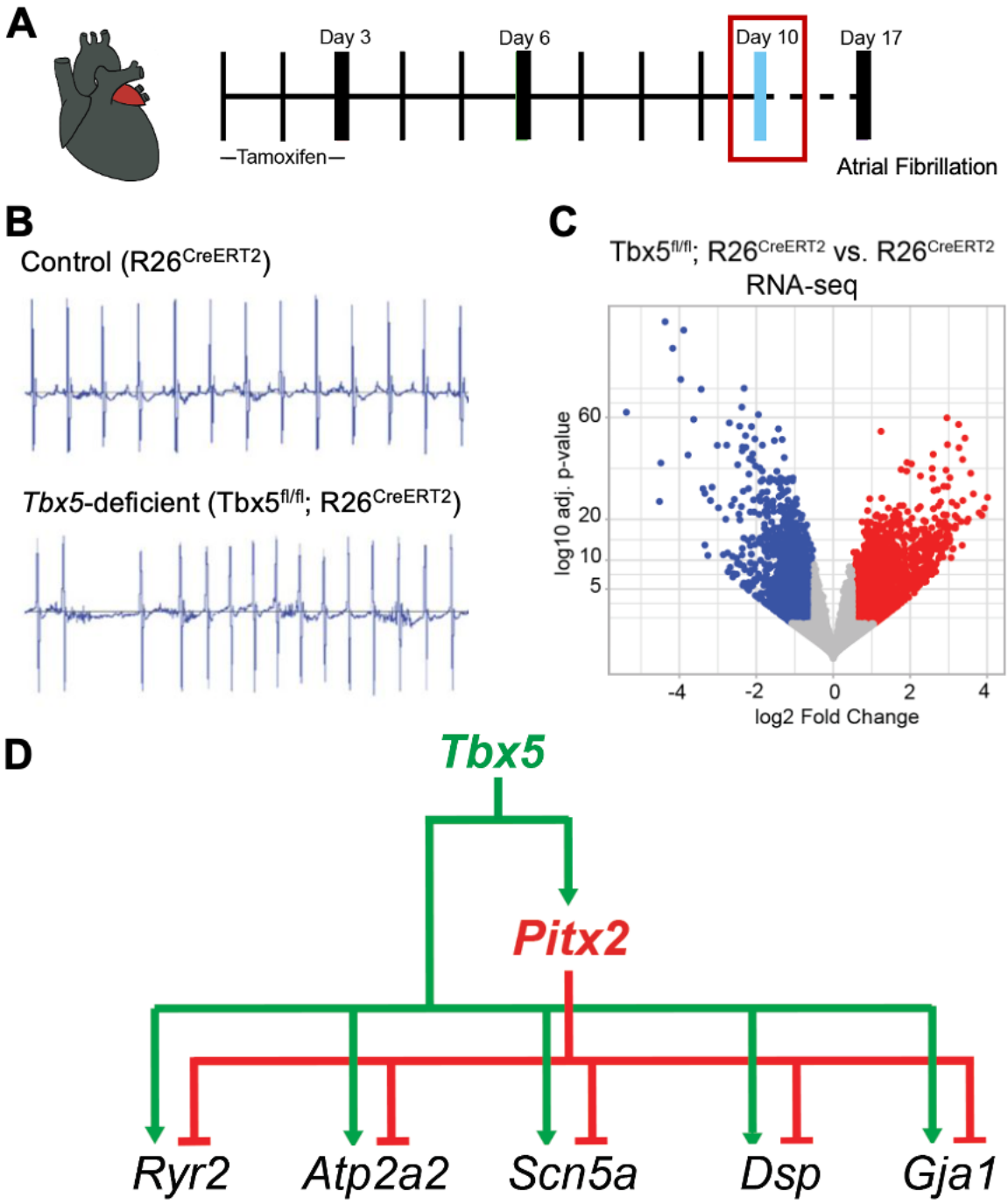


Figure 1.1. Removal of *Tbx5* in the adult mouse leads to atrial fibrillation, previously published by our laboratory (Nadadur et al., 2016) [118]. A. Schematic depicting the timeline of left atrial tissue collection 10 days after initiation of tamoxifen treatment. B. Representative ECG of $R26^{CreERT2}$ (control) and $Tbx5^{fl/fl}$;$R26^{CreERT2}$ (*Tbx5*-deficient) mice after tamoxifen regiment. C. Volcano plot of RNA-seq comparing transcriptional changes in the left atria of control and *Tbx5*-deficient mice. D. Schematic describing the incoherent feed-forward loop of cardiac rhythm genes regulated by TBX5 and PITX2.

Comorbidity of atrial fibrillation and heart failure

Atrial fibrillation and heart failure have a strong epidemiologic link due to the co-existence and predisposition to each other [121], [122]. Heart failure can lead to atrial fibrillation, and atrial fibrillation can lead to heart failure. Both cardiac diseases have an increased risk for death, but the outcome and prognosis for an individual with both is worse. Their co-existence has traditionally been attributed to shared risk factors including age, coronary artery disease, myocardial infarction, high blood pressure, and obesity. Interestingly, both atrial fibrillation and heart failure also undergo similar cardiac remodeling process including hypertrophy, fibrosis, inflammation, and electrophysiological changes. While there have been many comparisons of the epidemiologic link between heart failure and atrial fibrillation, the comparison of their gene regulatory networks may provide insight into the pathophysiology behind the comorbidity of atrial fibrillation and heart failure.

Heart failure is a debilitating disease that is described by the inability of the heart to pump a sufficient amount of blood to meet the demands of its workload. When the body is not getting an adequate amount of nutrient-rich blood the heart has compensatory processes that are initially crucial in maintaining heart functionality, but these changes are maladaptive and will eventually lead to heart failure. Heart failure is categorized into two distinct sub-types, heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction (HFpEF) [123], [124]. Ejection fraction is a measurement of the amount of blood that is pumped out of the left ventricle, and decreased ejection fraction measurements are an indication of heart malfunctions. HFrEF, also known as systolic heart failure, is characterized by left ventricular dilatation, ineffective systole to pump blood out of the heart. HFpEF, also known as diastolic heart failure, is characterized by left ventricular stiffness and muscle thickening, which leads to

ineffective diastolic relaxation. While there are distinctions between them, it is also clear that there are similarities including shared pathological remodeling processes as a result of insufficient amounts of blood being pumped from the heart. This maladaptive cardiac remodeling processes involve diverse cell types in the heart, and includes cardiomyocyte hypertrophy, apoptosis, electrophysiological alterations, fibrosis, and inflammation. Heart failure is estimated to affect 23-36 million individuals around the world, making it a major public health issue [125], [126]. In the United States alone there are around 6 million people with heart failure, but the prevalence of heart failure is projected to increase to more than 8 million people by 2030 [121], [127]. These staggering statistics warrant a better understanding of the molecular mechanisms driving the progression of this disease.

Fortunately many mouse models have been developed to interrogate the pathophysiology of heart failure [128]–[130]. Our laboratory utilized a transverse aortic constriction (TAC) mouse model provided by our collaborator David Park (NYU) [131], [132]. This is a widely used heart failure model that results in pressure overload-induced hypertrophy and heart failure in the mouse. TAC banding is accomplished by placing a tight suture around the aortic arch to make it difficult for the left ventricle to pump a sufficient amount of blood. The amount of constriction is determined by a needle placed alongside the aorta, in order to maintain consistency among the replicates. The gauge size of the needle will determine the level of constriction, which will also determine the severity of the phenotype [133]. The reproducibility among the replicates can also be maintained by quantifying the flow velocity to determine the pressure gradient. TAC banding has developed into a reliable surgical method to induce hypertrophy and heart failure in the mouse. While TAC is a dependable method used to induce cardiac pathology, the onset of the pressure-overload is instantaneous with the placement of the suture. This acute change in pressure is very

different from the gradual onset of hypertension seen in humans. Even with this caveat, TAC experiments have been valuable in identifying genes and pathways important in the pathophysiology of cardiac disease [132], [134]–[137]. In collaboration with David Park, our laboratory will utilize this heart failure mouse model to identify the shared regulatory mechanisms of atrial fibrillation and heart failure, two cardiac diseases with a strong epidemiologic link.

Cardiac remodeling in heart disease

The heart is composed of various cell types that work collaboratively, but cardiac perturbations can result in electrophysiologic, structural, or signaling changes in the heart (Figure 1.2). Many of these changes are thought to be acutely beneficial to compensate for the perturbation, however a prolonged exposure can lead them to become maladaptive and detrimental to normal heart function. Cardiomyocytes are the muscle cells behind the contractile force in the heart, and they are a terminally differentiated cell type. The cardiomyocytes will either undergo apoptosis or increase in size during cardiac stress. During hypertrophy, cardiomyocytes increase in size when the heart requires more force, which functions as a compensatory mechanism to meet the metabolic demands [123], [138]. There are several signaling pathways that can activate and regulate cardiac hypertrophy processes [123], [138]–[141]. For example, increases in intracellular calcium can activate calcineurin, which de-phosphorylates NFAT (nuclear factor of activated T cells) transcription factors; NFAT then travels to the nucleus to activate cardiac hypertrophy [138]–[140], [142]. Another cell type in the heart are the fibroblasts, which provide structural support to the heart by producing extracellular matrix proteins, including collagens. During cardiac stress the fibroblasts are activated and begin to excrete an excessive amount of extracellular matrix proteins [123], [143], [144]. These activated fibroblasts, or myofibroblasts, form fibrotic scars that can affect the heart's ability to function properly [123], [143]. Cardiac myofibroblasts originate from

resident fibroblasts, endothelium, epicardium, and bone marrow progenitors. Cardiac fibroblast activation is attributed to several different signaling pathways, including transforming growth factor β (TGF-β), angiotensin II (Ang II), and endothelin (ET-1) [123], [143], [144]. TGF-β signaling in particular is a well-studied canonical stimulus for fibroblast activation. The activation of this signaling pathway results in the phosphorylation of Smad2/3, binding to Smad4, which then travels to the nucleus to activate downstream gene expression of profibrotic genes [123], [144]–[146]. Fibroblasts in an *in vitro* setting can be treated with TGF-β in order to explore the transcriptional, molecular, and structural changes that occur during myofibroblast activation. Cardiomyocytes and fibroblasts are the most enriched cell populations in the heart, and we aim to identify the molecular mechanisms driving the pathophysiology in cardiac disease.

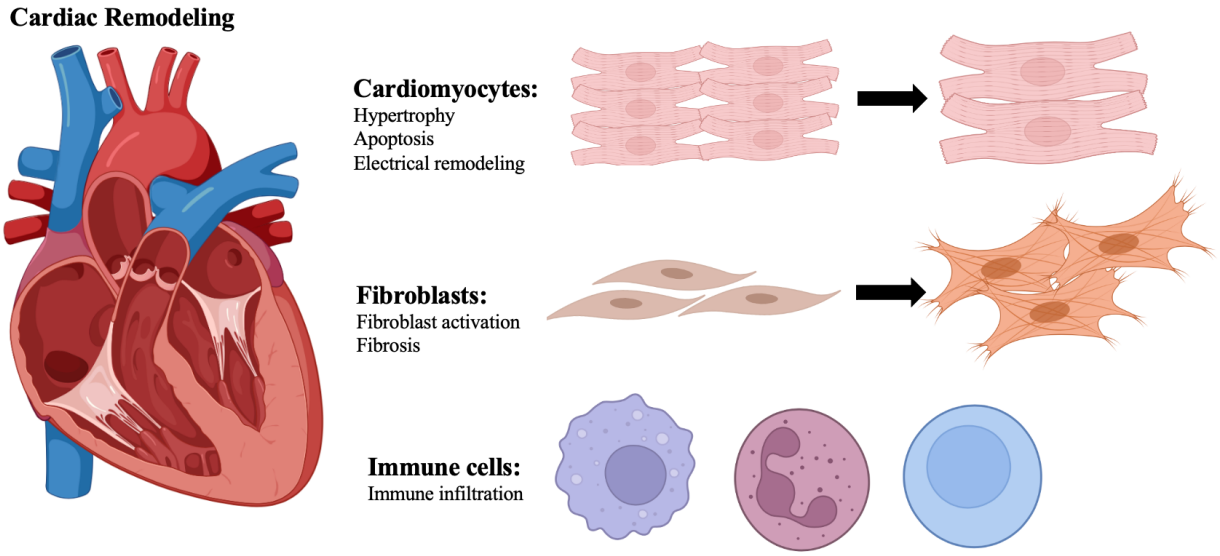


Figure 1.2. Schematic depicting the various cell types in the heart that undergo changes during cardiac remodeling processes.

Regulatory mechanisms of gene regulation

Homeostasis in the heart is maintained by very precise gene expression levels, to ensure

proteins are produced in the correct cell at the right time. Cardiac rhythm in particular is a very dose sensitive physiologic process, and any perturbations to this gene regulatory network can have severe consequences. The deletion and overexpression of cardiac rhythm genes can lead to disruptions to this fine-tuned network, and result in cardiac rhythm disorders [118], [147], [148]. The dose sensitive nature of this system is best exemplified in the dynamics between TBX5 and PITX2, two transcription factors that work antagonistically to control the expression of cardiac channel gene expression. The gene expression and AF inducibility seen in the TBX5 haploinsufficiency was rescued with PITX2 haploinsufficiency, providing evidence for the dose-dependent nature of gene expression important for cardiac rhythm. This is an example of transcription factors having a significant effect on the phenotype through to their regulation of downstream gene expression, but there are many alternative mechanisms by which gene expression is regulated.

An additional mechanism for titrating accurate gene expression is the utilization of miRNAs, which are small non-coding transcripts that post-transcriptionally repress gene expression [149]–[156]. miRNAs are transcribed and processed to form a pre-miRNA, which is exported to the cytoplasm for further processing until the final 22-bp non-coding transcript is complete. Several miRNAs can also be transcribed as one transcript if they are clustered together, and in many instances they regulate different target genes in a similar biological pathway. miRNAs can also affect the expression of several target genes, many of which are involved in the same processes. The specificity of the interaction between the miRNA and its target genes lies in the seed region of the miRNA and the 3'UTR of the mRNA transcript. miRNAs post-transcriptionally repress gene expression by inhibiting translation or marking the mRNA transcript for degradation [149]–[156]. The miRNAs are a feedback regulatory mechanism to fine tune gene expression and

act as a safeguard to maintain stable gene expression. miRNAs play an important role in many aspects of cardiovascular homeostasis and disease [154], [155]. Therefore, further identifying miRNAs and elucidating their downstream targets will refine our understanding of the gene regulatory mechanisms in place to maintain stable gene expression.

The transcriptional basis of homeostasis and disease is implied by the non-coding nature of genetic variation identified by GWAS, as the majority of GWAS signals lie within the non-coding regions of the genome [157]. Functional genomic approaches to unveil the gene regulatory networks relevant to disease is therefore a high priority. These non-coding regions contain enhancers that are bound by transcription factors to modulate gene expression. These enhancers can also be transcribed to produce non-coding RNA (ncRNA) transcripts [158]–[164]. Therefore, differential deep sequencing of ncRNA transcripts can be used to identify active enhancers in a context dependent manner, and also provide a quantitative measurement for enhancer activity [158], [159], [162]. For example, our laboratory has previously identified enhancers for important cardiac rhythm genes by comparing the ncRNA transcripts from control and *Tbx5*-deletion mice, and in the *Ryr2* locus the enhancer with the highest activity was adjacent to the ncRNA. The transcriptional profiling of the ncRNA transcripts can identify regulatory elements to provide insight into the gene regulatory network driving the expression of genes important for homeostasis and disease. The ncRNA profiling can also identify ncRNA transcripts that themselves have functional roles in the regulatory mechanisms of gene expression [159], [165]. Long non-coding RNA (lncRNA) transcripts in particular have been shown to be important in various aspects of cardiac disease, including hypertrophy and fibrosis [166]–[170]. MALAT1 and *Wisper* are lncRNAs that have increased expression following myocardial infarction, and the *in vivo* knockdown of MALAT1 attenuated cardiac fibrosis in the infarcted hearts [167], [168]. *Chast* is

a lncRNA that has increased expression during pressure overload, and inhibition of the transcript prevents cardiac remodeling in the mouse heart [170]. These are examples of lncRNA transcripts that are activated in a disease context, and their expression directly affects cardiac remodeling processes. In summary, transcriptional profiling of the non-coding genome will identify regulatory elements and potential ncRNA transcripts with functional roles to provide insights into the gene regulatory networks of disease.

Genomic profiling experiments provide insights into the gene regulatory networks that drive healthy and disease conditions. This body of work is aimed at elucidating the regulatory mechanisms in atrial fibrillation, identifying similarities of atrial fibrillation and heart failure, and how the transcriptional changes evolve through disease progression in atrial fibrillation. Each of these chapters is distinct in the mechanisms we utilize to unveil the molecular processes of cardiac disease.

.

CHAPTER 2: DEFINING THE TBX5-DEPENDENT miRNA REGULATORY NETWORKS OF ATRIAL FIBRILLATION

Chapter 2.1 Introduction

Normal cardiac conduction depends on coordinated electrical activity in the heart, and when this synchronization is disrupted it leads to cardiac arrhythmias that can have detrimental consequences. The most common arrhythmia is atrial fibrillation (AF), and it is characterized by random cardiac electrical activity that leads to an irregular heartbeat. AF is a growing epidemic that has been estimated to affect over 30 million individuals world-wide, and its prevalence is predicted to continue to increase [13], [21]. Individuals with AF have an increased risk of strokes and heart failure, which contributes to the healthcare burden caused by this disease.

While atrial fibrillation can result as a secondary effect to other cardiac diseases, it is becoming evident that there is also a heritable component to the development of the disease. Genome-wide association studies (GWAS) have associated genetic variants to AF risk in many ion channels, cardiac structural genes, and canonical cardiac transcription factors, including TBX5 [20], [21], [23], [26]–[31]. In addition to GWAS, a TBX5 gain-of-function mutation was identified in a familial study of Holt-Oram syndrome with skeletal abnormalities and paroxysmal atrial fibrillation [29], [76], [84]. TBX5 is a T-Box cardiac transcription factor important for heart development and is required for the conduction system in the adult mouse [82], [116], [117]. Our laboratory has since shown that the removal of *Tbx5* from the adult mouse leads to spontaneous, sustained atrial fibrillation [118]. *Tbx5*-deficient mice have an irregular heartbeat, prolonged action potential, abnormal depolarizations, and lack a p-wave; all of which are canonical characteristics of atrial fibrillation. To understand the molecular mechanisms of the atrial fibrillation phenotype we observe in the *Tbx5*-deletion mouse model, transcriptional profiling revealed expression

changes in many important genes for atrial rhythm including ion channels and cardiac structure genes. In particular, transcriptional profiling revealed that TBX5 directly regulates the expression of calcium handling genes, such as *Atp2a2*, *Ryr2*, and *Sln,* and their change in gene expression contributes to the atrial fibrillation phenotype. This application of transcriptional profiling revealed an incoherent feed forward loop with TBX5 driving expression of these genes, while PITX2 antagonistically repressed their expression [118]. Interestingly, the gene expression changes and AF inducibility seen in TBX5 haploinsufficiency were rescued with PITX2 haploinsufficiency. Therefore, the antagonistic relationships of TBX5 and PITX2 provides evidence that the downstream target genes are controlled in a dose-dependent manner [118].

The homeostasis of many physiologic processes, including cardiac rhythm, are dose-sensitive, thus making stable gene expression critical for proper function. Precise gene regulation is dependent on feedback mechanisms for accurate gene expression. A canonical feedback mechanism for steady gene expression is the use of miRNAs in regulatory networks. miRNAs are small, non-coding RNAs that function by post-transcriptionally suppressing gene expression [149]–[156]. miRNAs attenuate gene expression by marking the mRNA target transcript for degradation or inhibiting translation. The biogenesis of miRNAs initiates with the transcription of a primary RNA (pri-RNA) transcript that is processed with Drosha (RNase III) to form the pre-miRNA. The pre-miRNA is subsequently exported out of the nucleus with exportin-5, and further processing with Dicer (RNase III) finally produces the 22 bp, mature miRNA transcript. miRNAs generally bind to the 3' UTR of target transcripts, and the specificity of this interaction is established between the complementarity of the mRNA and the miRNA seed region.

miRNAs have been shown to have essential roles in the every aspect of development and disease, including in the heart. miR-1 is an example of a well-studied miRNA important during

cardiac development and in the adult heart for proper cardiac function [154], [155], [171]. miR-1 is clustered near miR-133 and they are transcribed together prior to downstream processes that generate the two mature miRNA transcripts [171]. Interestingly, these miRNAs display functional redundancy, as transcription from two separate gene loci produce the same mature miRNA. For example, mature transcripts for miR-1 are transcribed from two distinct gene loci, to generate miR-1-1 and miR-1-2. The functionality of miRNAs can be evaluated through deletion and overexpression studies. The absence of miR-1-2 halves the amount of mature miR-1, which results in partial embryonic lethality from ventricular septal defects due to abnormal levels of *Hand2* in the embryo [172], [173]. The mice that did survive postnatally from the loss of miR-1-2 exhibited electrophysiologic defects including a shortened PR interval and prolonged ventricular depolarization, which was attributed to increased levels of *Irx5*, a regulator of potassium channel gene expression and another direct target of miR-1 [172]. miRNAs can control the expression of several downstream proteins, as it has also been shown that miR-1 targets KCNJ2 (potassium channel) and GJA1 (gap junction connexin), cementing its role as a critical miRNA for cardiac gene regulation [174]. miR-1 also has implications for human disease, as miR-1 levels are altered in various cardiac perturbations and has been examined as a putative biomarker [175]–[178]. These studies provide an example of a miRNA that plays an important role in electrophysiology, and the detrimental effects that the inappropriate dose of the miRNA and its downstream targets can have on proper cardiac rhythm.

Cardiac rhythm is a highly-dose sensitive biological process, and miRNAs provide a feedback mechanism to fine tune gene expression post-transcriptionally. While the importance of miRNAs have been implicated in various cardiac processes, including atrial rhythm, we wanted to explore the changes in miRNA expression in the context of our atrial fibrillation mouse model. To

evaluate the potential miRNA candidates that affect the pathophysiology of atrial fibrillation, we transcriptionally profiled small RNAs from the left atria of control and *Tbx5*-deletion mice.

The differentially expressed small RNAs identified in this experiment nominate putative miRNAs important in the TBX5-dependent gene regulatory network, with the potential to identify miRNAs instrumental in cardiac rhythm homeostasis. Several of the miRNAs identified in the miRNA-seq were subsequently evaluated in a high-throughput screen, and the overexpression of several candidate miRNAs resulted in arrhythmogenic phenotypes. This approach allowed us to examine several candidate miRNAs simultaneously, and also provide a model for arrhythmias in an *in vitro* setting in a dish. We focused on one of the most significantly down-regulated miRNAs in the miRNA-seq, miR-10b, which also displayed arrhythmogenic properties in the high-throughput screen. This is a particularly interesting candidate because miR-10b has been shown to repress the expression of TBX5 by targeting the 3' UTR region of the transcript [179]. Whole mouse and cellular electrophysiology studies revealed miR-10b mis-regulation leads to atrial fibrillation susceptibility. Evidence of a reciprocal relationship between miR-10b and TBX5 supports the notion that miR-10b is an important regulator of cardiac rhythm. Understanding the role of miRNAs in the TBX5-dependent gene regulatory network provides more insight into the regulatory mechanisms of cardiac rhythm and atrial fibrillation.

Chapter 2.2 Materials and Methods

Mouse model for atrial fibrillation

The conditional *Tbx5* deletion was completed with a floxed allele, and was previously described by our laboratory [64], [118], [159]. The third exon of *Tbx5* was excised with a tamoxifen inducible Cre recombinase driven by the R26, a locus used for stable, general gene expression. $Tbx5^{fl/fl}$;$R26^{CreERT2}$ (*Tbx5*-deletion) and $R26^{CreERT2}$ (control) mice were subjected to a tamoxifen regiment at 6-10 weeks of age as previously described [118]. The mice used in these experiments were raised in accordance with the Guide for the Care and Use of Laboratory Animals. All experiments were approved by the University of Chicago Institutional Animal Care and Use Committee (IACUC).

Small RNA transcriptional profiling

Left atria was isolated from $Tbx5^{fl/fl}$;$R26^{CreERT2}$ (*Tbx5*-deletion) and $R26^{CreERT2}$ (control) mice. RNA was prepared with a TRIzol-based extraction and isolated using a Qiagen RNeasy Column Kit. The mirVana miRNA Isolation Kit (Invitrogen) was used to isolate the miRNA transcripts. Library preparation and sequencing was performed by the University of Chicago Genomics Core Facility. The small RNAs were sequenced 10-20 million reads per sample. The reads were mapped to UCSC *Mus musculus* genome (mm9) using Tophat2, and differential gene expression was evaluated with Cuffdiff and miRbase.

High-throughput screen of miRNA candidates

The high-throughput kinetic imaging was completed by our collaborators in the laboratory of Alexandre Colas at Sanford Burnham Prebys. The high-throughput experiment is a screen to measure calcium transitions in cardiomyocytes transfected with miRNAs. Candidate miRNA were

identified in the small RNA-seq, and transfected in atrial-like cardiomyocytes at day 25. Three days after transfection the cells were stained with voltage-sensitive probes. A high speed microscope was used to record 5 second movies at 100 Hz and 20x magnification. The changes in fluorescence were measured over time and processed to calculate the action potential traces. Physiologic measurements were assessed from the action potentials and compared to the control cells; the measurements include APD75 and arrhythmia index (AI). The APD75 (action potential duration 75) measurement is the amount of time it takes to get to 75% repolarization. The APD75 is calculated from the median measurement within a cell, median within the transfected well, and then calculating the mean across the replicates. The arrhythmia index is a calculation designed to measure the variability in beat-to-beat intervals. The standard deviation of the distance from one peak to the next peak is divided by the mean of that distance, multiplied by 100. This calculation is measured within one cell, followed by the median of cells within each well, and then the mean of different replicates is finally calculated.

Deletion of miR-10b in the mouse

A floxed allele of miR-10b was used to generate a conditional deletion of miRNA in the adult mouse. To generate a cardiomyocyte-specific deletion of miR-10b, a tamoxifen inducible Cre recombinase for *Myh6* was used for the experiment. The *miR-10b*$^{fl/fl}$; *Myh6*^{CreERT2} mice were injected with tamoxifen (*miR-10b*-deletion) or corn oil (control). Tamoxifen regiment was initiated in mice 6-10 weeks of age.

Intracardiac electrophysiology studies

Catheter-based intracardiac recordings were completed 10-weeks after receiving tamoxifen. Mice were anesthetized with isoflurane, and a 1.1-F octapolar catheter (EPR-800, Millar Instruments) was inserted via the right jugular vein, and it was used to record right atrial

and ventricular electrograms. Programmed extra-stimulation protocols and burst pacing were used in an attempt to induce atrial fibrillation and atrial tachycardia.

Cellular electrophysiological recordings

A whole-cell patch-clamp method was used to record action potentials. Langendorff perfusion with collagenase type 2 (Worthington Biochemical) was used to isolate atrial myocytes, which were plated and incubated on laminin-coated glass dishes for 30 minutes. Atrial cardiomyocytes were perfused with Tyrode's solution during recordings. An Axopatch-200B amplifier connected to a Digidata1550A acquisition system (Axon Instruments) was used to record the whole cell action potentials. Action potentials were triggered using 0.5 nA x 2-ms current clamp pulses following liquid junction potential correction. A frequency of 2 kHz was used to filter the recordings using a built-in Bessel filter and sampled at 10 kHz. Results were analyzed using pCLAMP10 (Axon Instruments).

miR-10b mimic and inhibitors

HL-1 cardiomyocytes were transfected with miR-10a/b mimics and inhibitors (Thermo Scientific). The cells were co-transfected with mCherry, and 48-96 hours post transfection the cells were sorted for mCherry. Relative *Tbx5* expression levels were measured using qPCR.

TBX5 ChIP in HL-1 cardiomyocytes

HL-1 cardiomyocytes were collected and chromatin immunoprecipitation (ChIP) was performed using a Santa Cruz Biotechnology TBX5 antibody (sc-17866), as described in previous publications from our laboratory [118], [180]. Tissues were fixed in formaldehyde, washed, and then homogenized. Sonication was performed to shear the chromatin to approximately 100-300 bp. Samples were incubated overnight with TBX5 antibody bound to magnetic Dynabeads (Invitrogen). Beads were precipitated, washed, and the chromatin was eluted. When the samples

were de-crosslinked, the chromatin was purified using a Qiagen PCR cleanup kit. qPCR was used to determine fold enrichment between input controls and DNA bound to immunoprecipitation complex.

Chapter 2.3 Results

Small RNA-seq profiling identifies TBX5-dependent miRNAs.

Our laboratory has previously demonstrated that the conditional deletion of TBX5 from the adult mouse resulted in spontaneous, sustained atrial fibrillation [118]. TBX5 expression is instrumental for cardiac rhythm, as it is an important regulator of calcium-handling and ion channel gene expression. Calcium cycling is a crucial player in the excitation-contraction processes of the cardiomyocyte, therefore tightly controlled gene expression is essential for proper cardiomyocyte function. miRNAs are small non-coding RNAs that provide dose-sensitive feedback mechanisms to regulate gene expression, therefore our subsequent studies are aimed at identifying TBX5-dependent miRNAs and interrogating their putative function in the preservation of proper cardiac rhythm. Conditional deletion of *Tbx5* was accomplished with a tamoxifen inducible Cre recombinase at the Rosa26 locus in conjunction with a *Tbx5* floxed allele ($Tbx5^{fl/fl}$;$R26^{CreERT2}$). The mice were treated with tamoxifen for 3 days at 6-8 weeks of age. The left atria of $Tbx5^{fl/fl}$;$R26^{CreERT2}$ and $R26^{CreERT2}$ mice were collected one week after tamoxifen treatment for small RNA profiling. The small RNAs from these tissues were isolated for sequencing of the miRNAs. A comparison of the control and *Tbx5*-deficient atria showed distinct expression patterns between the two genotypes, and consistent expression among the biological replicates (Figure 2.1A). The removal of *Tbx5* resulted in 162 differentially expressed miRNA candidates, including members of the miR-10 and miR-154 family of miRNAs. All 7 members of the miR-10 gene family were down-regulated following the removal of *Tbx5*, while 8 of 13 miRNAs in the miR-154 gene family were up-regulated (Figure 2.1C-D). This sequencing experiment has identified *Tbx5*-dependent miRNAs that have changed expression in the absence of *Tbx5* and potential miRNAs regulating the downstream effects.

High-throughput screen identifies miRNA candidates that lead to an arrhythmogenic phenotype

The differentially expressed candidate miRNAs were identified prior to the onset of arrhythmogenic phenotypes in the mouse model, therefore these candidates are potentially involved in the regulation of cardiac rhythm gene expression. In an effort to narrow the scope of candidates to those that may be important for cardiac rhythm homeostasis, we assessed them in a high-throughput calcium screen performed by our collaborators in the lab of Alexandre Colas at Sanford Burnham Prebys. Atrial-like cardiomyocytes were transfected with miRNAs or anti-miRNAs, and subsequently stained with a fluorescent voltage sensitive probe and recorded with a high speed microscope (Figure 2.2A). The changes in fluorescence in the recordings were a measurement of the action potential in the cell. Representative tracings were shown for the control, miR-10a, miR-10b, and miR-449c (Figure 2.2B-E). The action potential duration (APD) was evaluated for the candidate miRNAs. It was evident that dependent on the transfected candidate miRNA, there is an elongation or shortening of the APD (Figure 2.2F-I, Figure 2.S1E). For example, the transfection of miR-10a and miR-10b resulted in prolongation of the APD, while the presence of miR-449c resulted in a shortened APD, as shown in the representative action potentials (Figure 2.2F-H). The prolongation and shortening of an APD can lead to triggered activity, which is a mechanism of cardiac arrhythmias. Therefore, we quantified a measurement of the irregularity between peak distances, termed the Arrhythmia Index (AI), which is used as a metric to identify miRNA candidates that have an effect on physiologic measurements (Figure 2.2J, Figure 2.S1F). This measurement identifies miRNA overexpression that leads to beat-to-beat irregularity, and acts as a proxy for an arrhythmogenic phenotype in the atrial-like cardiomyocytes. miR-10a, miR-10b, and miR-449c, which had shortened and prolonged APDs, also had an increased AI compared to the control (Figure 2.2J). A closer inspection of the action potential traces revealed the

cardiomyocytes transfected with miRNAs displayed a variety of arrhythmia-like phenotypes, including ectopic beats with the overexpression of miR-1231 and miR-210 (Figure 2.S1C-D). In summary, the high-throughput screen identified *Tbx5*-dependent miRNAs with a potential role in regulating cardiac rhythm, as their overexpression led to an arrhythmia-like phenotype.

To further explore the potential role of the miRNAs identified in the differential miRNA sequencing for the maintenance of cardiac rhythm, we focused our studies on miR-10b. Of the candidate miRNAs identified, miR-10b was one of the most highly expressed and mis-regulated miRNAs. The high-throughput screen also demonstrated overexpression of the miR-10b resulted in a prolonged action potential, a trigger for arrhythmogenic phenotypes. We confirmed this phenotype with cellular electrophysiology studies in HL-1 cardiomyocytes transfected with miR-10b mimics (Figure 2.S2). The miRNA was also of particular interest because previous studies have shown that miR-10b targets TBX5 to regulate its expression [179], [181]. We have previously demonstrated that atrial cardiomyocytes isolated from a *Tbx5* conditional deletion have calcium handling abnormalities, which leads to a prolonged action potential [118]. This is the same phenotype that is present when miR-10b was overexpressed.

Removal of miR-10b leads to atrial fibrillation susceptibility

Both loss and gain-of-function mutations in TBX5 have been implicated in atrial fibrillation and Holt-Oram syndrome [84], [182]. Precise gene expression is necessary for the maintenance of cardiac rhythm genes, therefore any alterations to this finely tuned system can have detrimental effects. Since TBX5 is a known target of miR-10b, we hypothesized that the alterations to miR-10b would also have an impact on atrial rhythm. Therefore we examined the role of miR-10b in an *in vivo* context using telemetry ECG recordings in cardiac specific miR-10b deletion mice (miR-10b$^{fl/fl}$, *Myh6*^{CreERT2}, injected with tamoxifen) compared to control mice (miR-10b$^{fl/fl}$,

Myh6^{CreERT2}, injected with corn oil). No observable phenotype was observed following the removal of miR-10b. Echocardiography was completed to evaluate cardiac function, and no change in cardiac function was detected after the removal of miR-10b. To further elucidate the role of miR-10b electrophysiology, pacing studies were performed on miR-10b deletion and control mice (Figure 2.3). Mice with cardiomyocyte-specific miR-10b deletion (miR-10b$^{fl/fl}$, *Myh6*^{CreERT2}, tamoxifen) were highly susceptible to atrial fibrillation, while none of the control mice (miR-10b$^{fl/fl}$, *Myh6*^{CreERT2}, corn oil) had cardiac rhythm abnormalities after pacing (Figure 2.3B-F). Thus, the cardiomyocyte-specific removal of miR-10b leads to increased susceptibility to atrial fibrillation. Physiologic measurement of the intracardiac electrogram reveals an elongation of the A-H interval, an indicator of AV node function (Figure 2.3G). The H-V interval measurement is an indicator of distal conduction, and there was no observable difference after the removal of miR-10b (Figure 2.3H). Overall, the removal of miR-10b from the cardiomyocytes in the adult mouse leads to atrial fibrillation susceptibility, reinforcing the importance of miRNAs in maintaining gene expression levels for stable cardiac rhythm.

Regulatory loop between TBX5 and miR-10b

We have previously shown the conditional deletion of TBX5 in the adult mouse leads to spontaneous, sustained atrial fibrillation. The deletion of *Tbx5* is a consequence of the removal of exon 3, resulting in a shortened transcript, and an inactive protein [64]. Following the deletion of *Tbx5*, the sequencing results perplexingly showed that *Tbx5* was upregulated (Figure 2.4A). A closer inspection confirmed the loss of exon 3, but intriguingly the remaining exons has increased expression (Figure 2.4B). A potential explanation for the upregulation of the inactive *Tbx5* transcript is changes to post-transcriptional repression due to miRNAs. MiR-10a and miR-10b were particularly appealing candidates because they have previously been shown to regulate the

expression of TBX5 by binding the 3' UTR to inhibit translation [179]. We have confirmed the ability of miR-10a and miR-10b to regulate the expression of *Tbx5*, through the transfection of mimics and inhibitors. The transfection of miR-10a or miR-10b mimics into HL-1 cardiomyocytes led to decreased expression of *Tbx5* transcript compared to the control, while the transfection of miR-10a/b inhibitors resulted in an increased expression of *Tbx5* (Figure 2.4C). Therefore, miR-10a and miR-10b are validated to regulate the expression of *Tbx5*. Interestingly, these miRNAs were also identified as differentially expressed in the left atria of the *Tbx5* conditional deletion mouse model. The downregulation of miR-10a and miR-10b could be a direct result from the deletion of *Tbx5* or a secondary effect due to this perturbation. To determine whether the changes to miRNA expression are due to TBX5, our laboratory performed chromatin immunoprecipitation quantitative polymerase chain reaction (ChIP-qPCR) for TBX5 in HL-1 cardiomyocytes. ChIP-qPCR established a 50-fold enrichment for TBX5 at the miR-10a and miR-10b promoter, suggesting a direct relationship between TBX5 and the expression of miR-10a/b (Figure 2.4D). Therefore, TBX5 drives expression of miR-10a/b and then the miRNAs negatively regulate the expression of *Tbx5*, which in turn helps to maintain a steady level of gene expression (Figure 2.4F). These studies provide insight into the regulatory mechanisms of *Tbx5* expression, and the ability of miRNAs to act as a buffer for stable transcription factor gene expression to ensure appropriate expression of downstream cardiac rhythm genes.

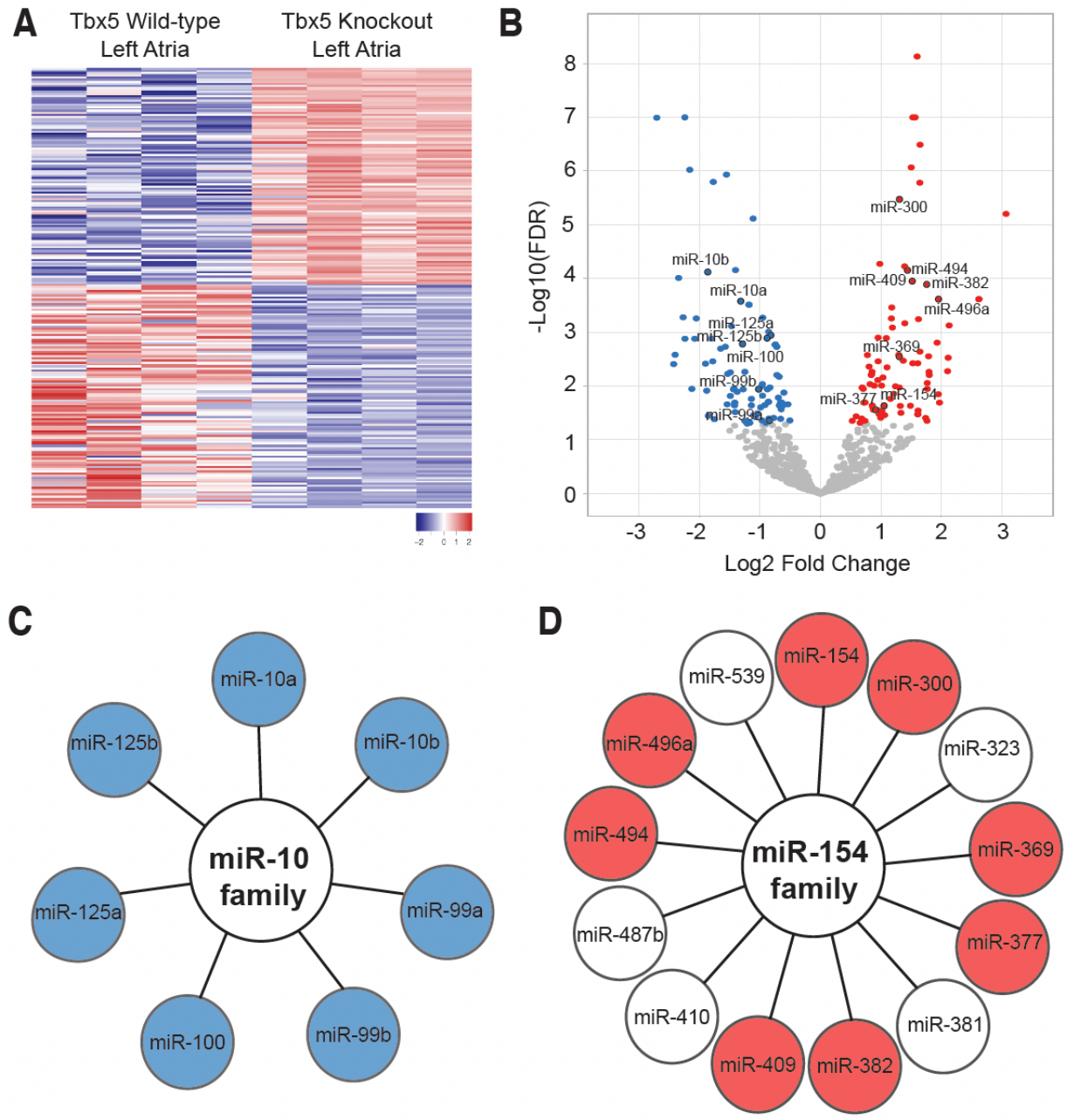


Figure 2.1. Identification of differentially expressed miRNAs in the *Tbx5*-deletion mouse model. A. Heatmap of the small RNA transcriptional profiling in the left atria of *Tbx5*-deletion (*Tbx5*$^{fl/fl}$; *R26*^{CreERT2}, n=4) and control (*R26*^{CreERT2}, n=4) mice displays a consistent change in expression among the replicates. B. Volcano plot of relative expression from the left atria of *R26*^{CreERT2} and *Tbx5*$^{fl/fl}$; *R26*^{CreERT2} mice shows 162 microRNAs with significantly altered expression (FDR<0.05 and |log2FC|>0.5); all downregulated miRNAs are labeled in blue, all upregulated miRNAs are labeled in red, and all nonsignificant miRNAs are labeled in grey. C. The miR-10 family of miRNAs were significantly downregulated in the *Tbx5*-deficient tissue and highlighted in blue: miR-10a (log2FC= -1.31, p= 0.0003), miR-10b (log2FC= -1.86, p=7.16E-5), miR-99a (log2FC=

Figure 2.1, continued

-0.84, p= 0.04), miR-99b (log2FC= -1.02, p= 0.01), miR-100 (log2FC= -1.29, p=0.002), miR-125a (log2FC= -0.82, p=0.001), and miR-125b (log2FC= -0.88, p=0.001). D. Several members of the miR-154 family of miRNAs were significantly upregulated in the *Tbx5*-deficient tissue and highlighted in red: miR-154 (log2FC= 1.05, p=0.02), miR-300 (log2FC= 1.74, p=0.04), miR-369 (log2FC=1.30, p=0.003), miR-377 (log2FC= 0.91, p=0.03), miR-382 (log2FC=1.76, p=0.0001), miR-409 (log2FC=1.52, p=0.0001), miR-494 (log2FC=1.44, p=6.97E-5), and miR-496a (log2FC=1.95, p=0.0002).

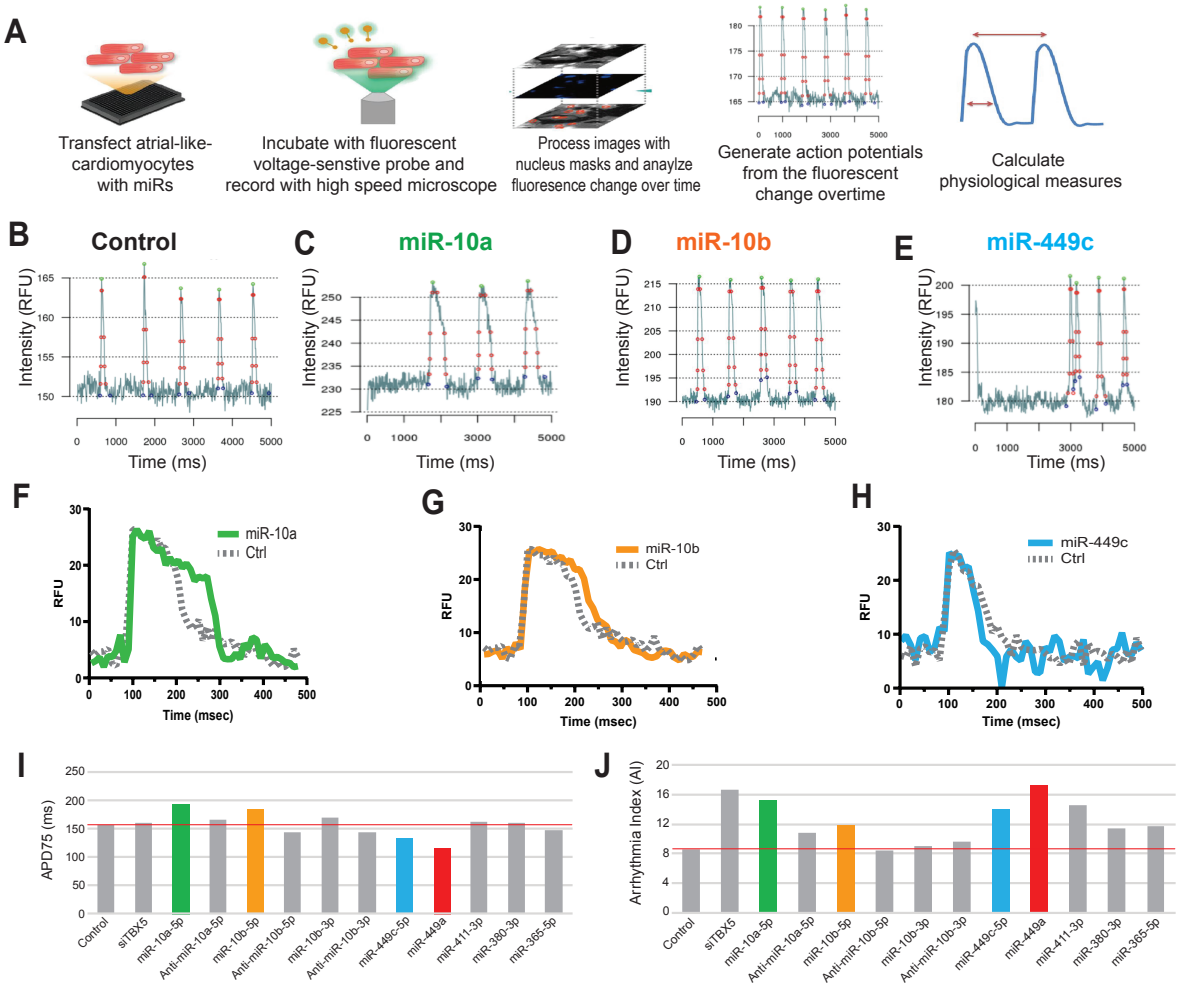


Figure 2.2. A high-throughput screen reveals arrhythmogenic phenotypes. A. Workflow of high-throughput calcium screen includes transfection of atrial-like cardiomyocytes (ACM) with miRNAs or siRNAs. Three days post-transfection, the cells are incubated with a fluorescent voltage-sensitive probe and recorded with a high-speed microscope. Then, any changes in fluorescence are translated to action potentials. Physiological measurements were calculated from the action potentials. B-E. Representative fluorescence measurements of the control (B) compared to cardiomyocytes transfected with miR-10a (C), miR-10b (D), miR-449c (E). F-H. Representative action potential of the control compared to cardiomyocytes transfected with miR-10a (F), miR-10b (G), and miR-449c (H), respectively. I. APD75 measurements for miRNA and siRNA transfections in cardiomyocytes. J. Arrhythmia Index (AI) for cardiomyocytes transfected with miRNAs.

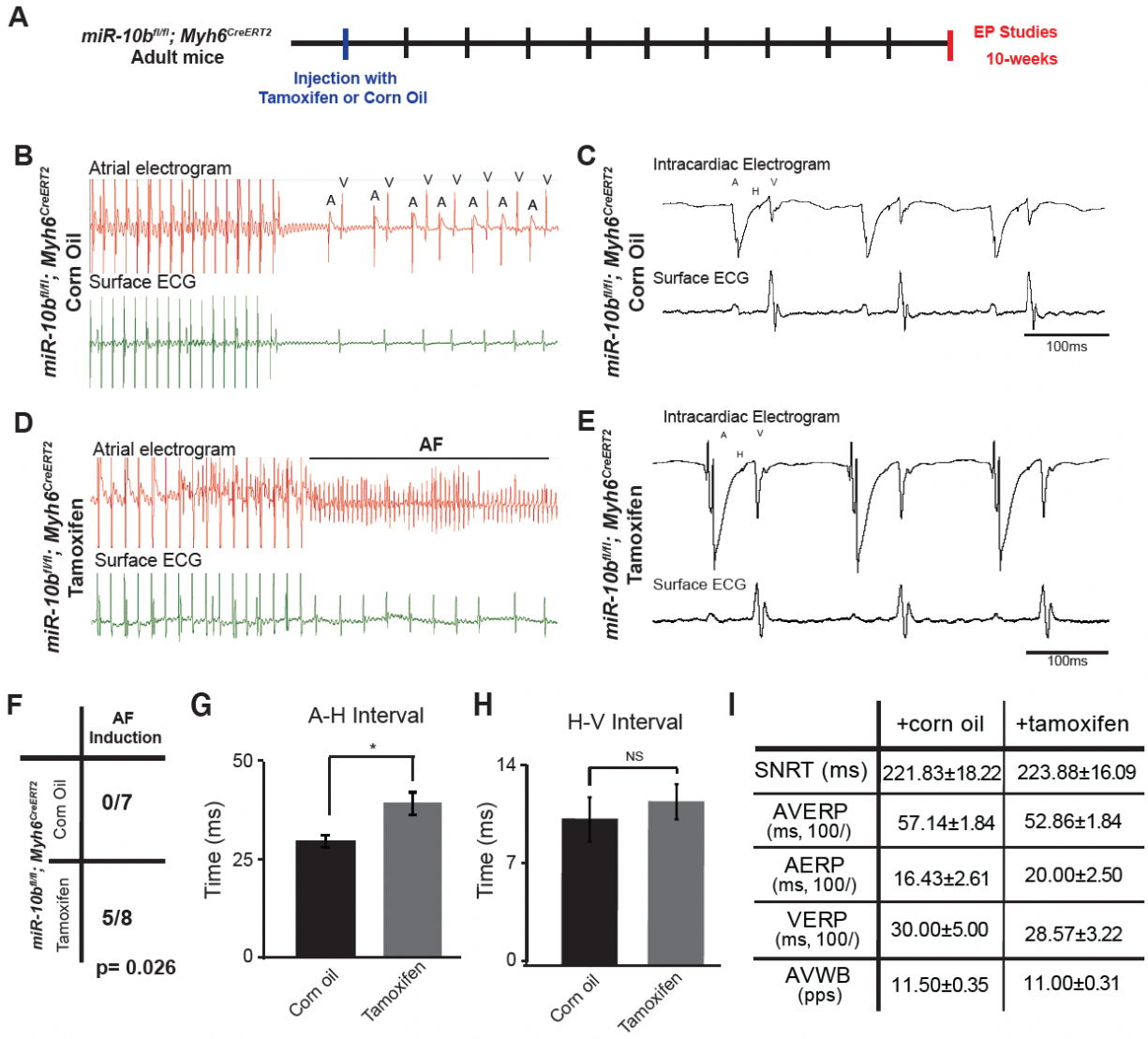


Figure 2.3. The depletion of miR-10b leads to atrial fibrillation susceptibility. A. Whole organ mouse electrophysiology workflow shows *miR-10b$^{fl/fl}$; Myh6^{CreERT2}* mice at 6-10 weeks of age injected with corn oil or tamoxifen, and rapid cardiac pacing 10 weeks post-injection. B-E. Atrial and surface ECGs of *miR-10b$^{fl/fl}$; Myh6^{CreERT2}* mice injected with corn oil (B-C) or tamoxifen (D-E). Corn oil injected mice display regular sinus rhythm post rapid atrial pacing (B). Tamoxifen injected mice reveals AF after rapid atrial pacing (D). F. A significant number of miR-10b deletion (*miR-10b$^{fl/fl}$; Myh6^{CreERT2}* tamoxifen injected; 5/8) mice were induced into AF through rapid atrial pacing, while no control (*miR-10b$^{fl/fl}$; Myh6^{CreERT2}* corn oil; 0/7) mice were induced into AF. G. A-H interval calculations from *miR-10b$^{fl/fl}$; Myh6^{CreERT2}* mice injected with corn oil or tamoxifen displayed a significant increase in A-H interval compared to the control. H. H-V interval from *miR-10b$^{fl/fl}$; Myh6^{CreERT2}* mice injected with corn oil or tamoxifen displayed no significant difference. I. Additional electrophysiologic measurements of the *miR-10b$^{fl/fl}$; Myh6^{CreERT2}* mice injected with corn oil or tamoxifen.

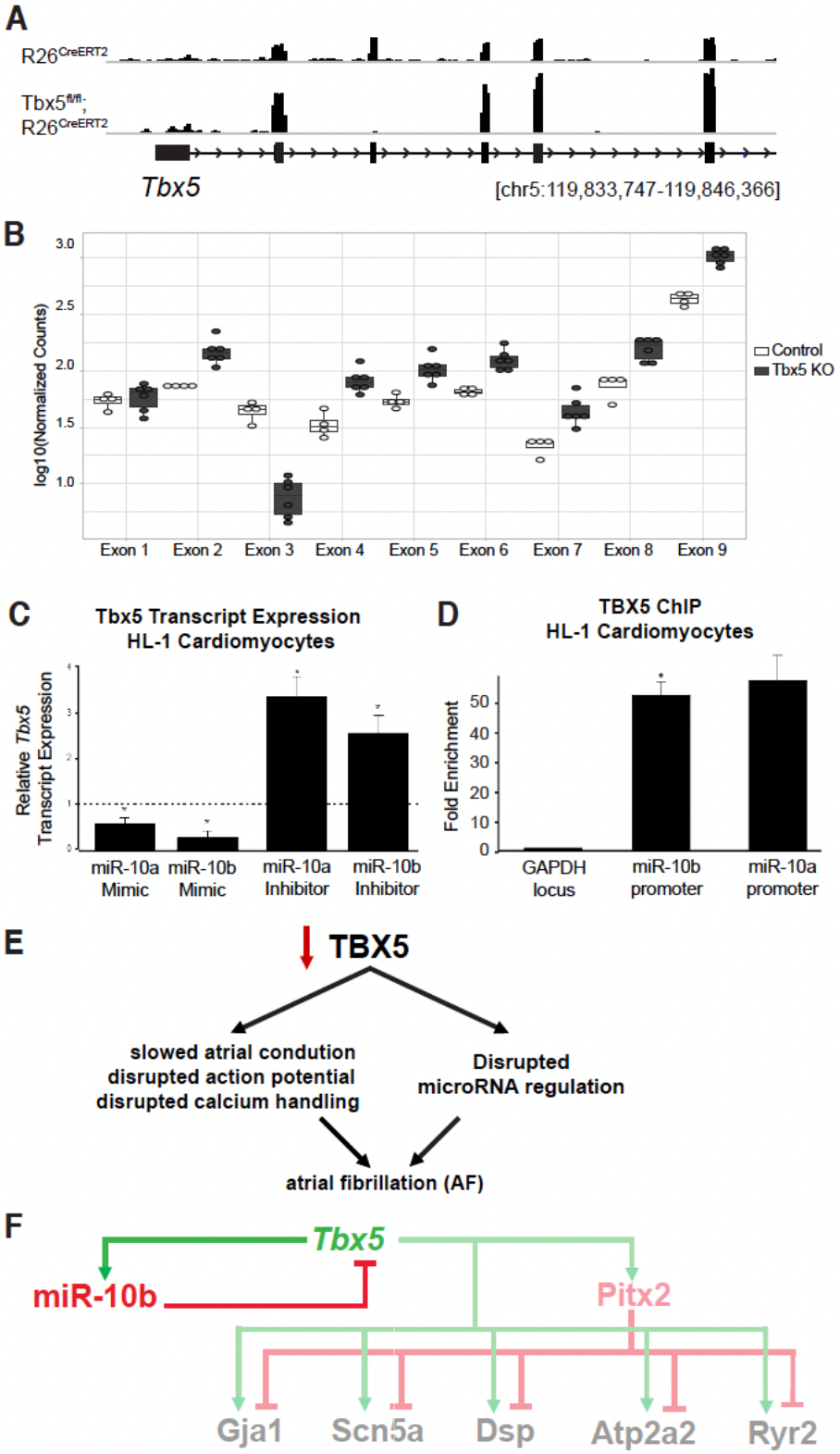


Figure 2.4. miR-10b is implicated in a TBX5-dependent network crucial in maintaining cardiac rhythm. A. *Tbx5* genomic locus [mm10, chr5:119,833,747-119,846,366] for the RNA-seq from the *Tbx5*-deletion (*Tbx5*$^{fl/fl}$; *R26*^{CreERT2}) and control (*R26*^{CreERT2}) mice. B. Quantification of the

Figure 2.4, continued

counts per exon in the *Tbx5*-deletion (*Tbx5*$^{fl/fl;}$ *R26*^{CreERT2}) and control (*R26*^{CreERT2}). C. HL-1 cardiomyocytes transfected with miR-10a or miR-10b mimics show reduced *Tbx5* expression. HL-1 cardiomyocytes transfected with miR-10a or miR-10b inhibitors reveal increased *Tbx5* expression. D. TBX5 ChIP in HL-1 cardiomyocytes reveals that TBX5 is highly enriched at the miR-10a and miR-10b promoters. E. Schematic depicting the relationship between TBX5 and atrial fibrillation, which includes a miRNA network. F. Schematic depicting the TBX5 gene regulatory network. The expression of *Tbx5* is repressed by miR-10b, confirmed through mimic and inhibitor experiments. The expression of miR-10b is regulated by TBX5, therefore providing a negative feedback loop.

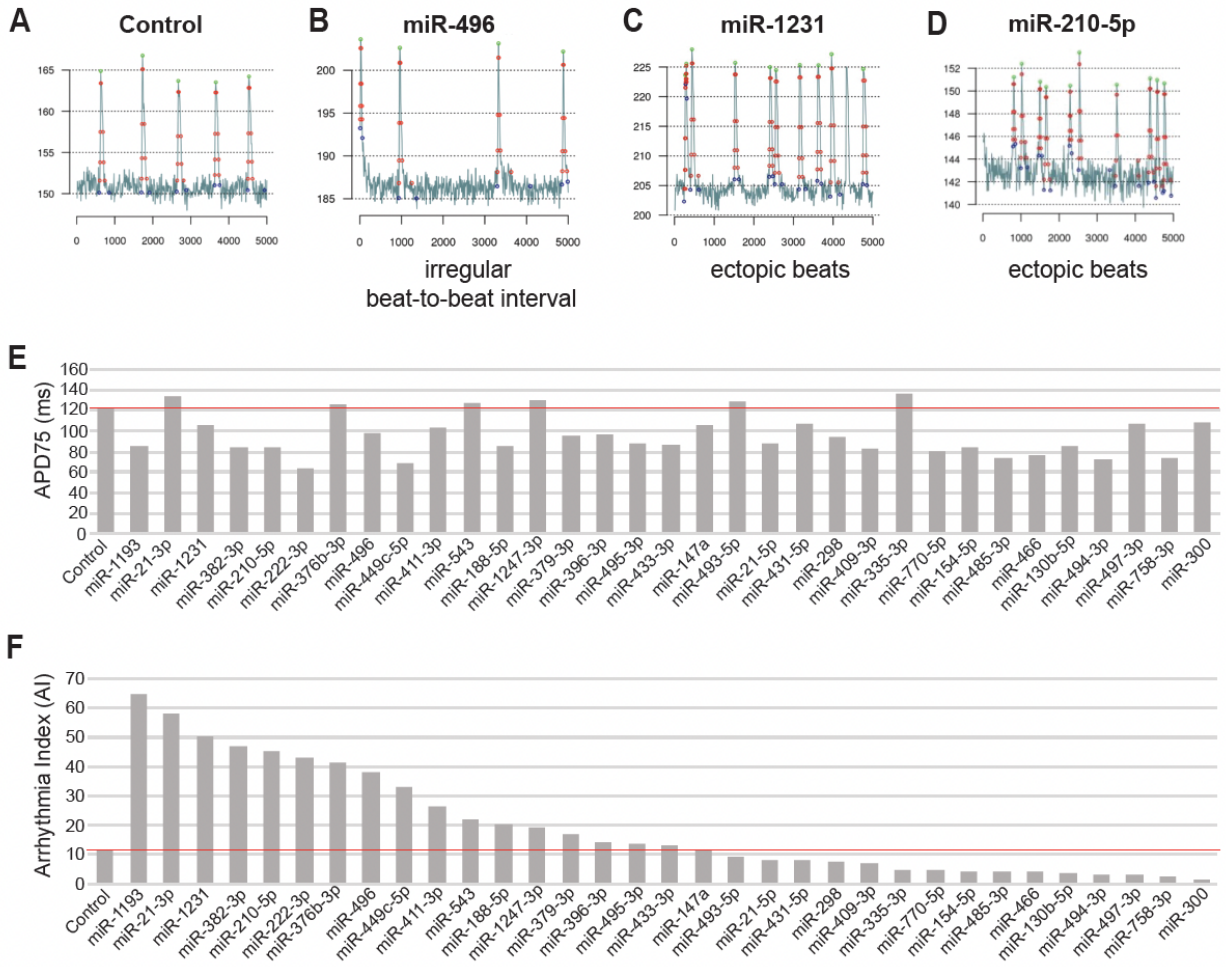


Figure 2.S1. Additional miRNAs tested in high-throughput screen. A-D. Representative action potential of the cardiomyocytes transfected with control (A), miR-496 (B), miR-1231(C), miR-210 (D). E. APD75 measurements for miRNA transfections in cardiomyocytes. F. Arrhythmia index (AI) for cardiomyocytes transfected with miRNAs.

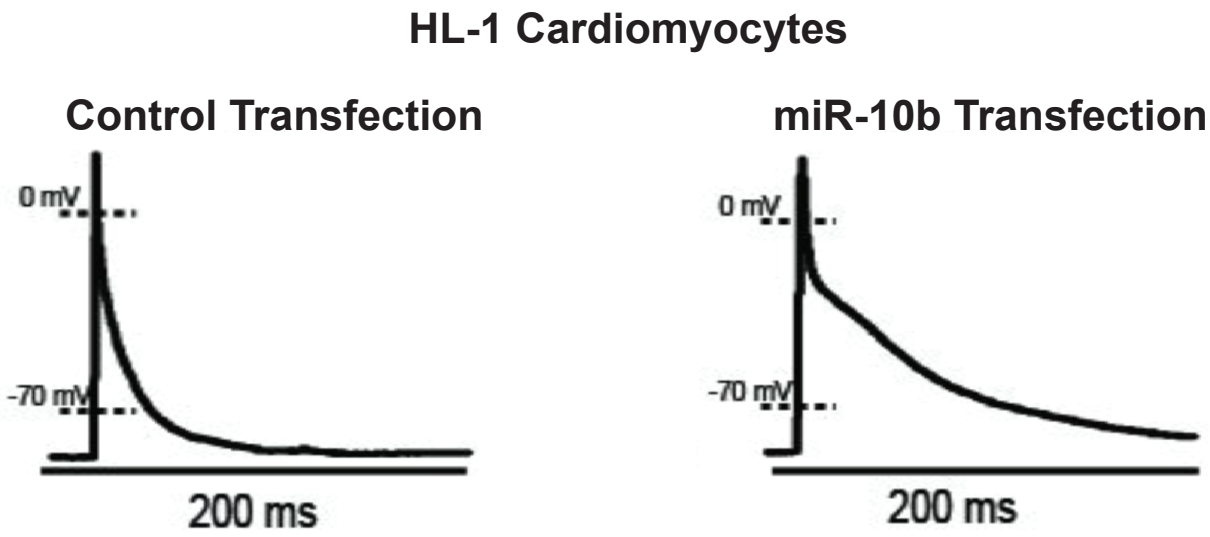


Figure 2.S2. Prolonged action potential after the transfection of miR-10b in cardiomyocytes. Representative action potentials of HL-1 cardiomyocytes transfected with a control (left) and miR-10b mimic (right). The miR-10b overexpression cardiomyocytes show an elongated action potential.

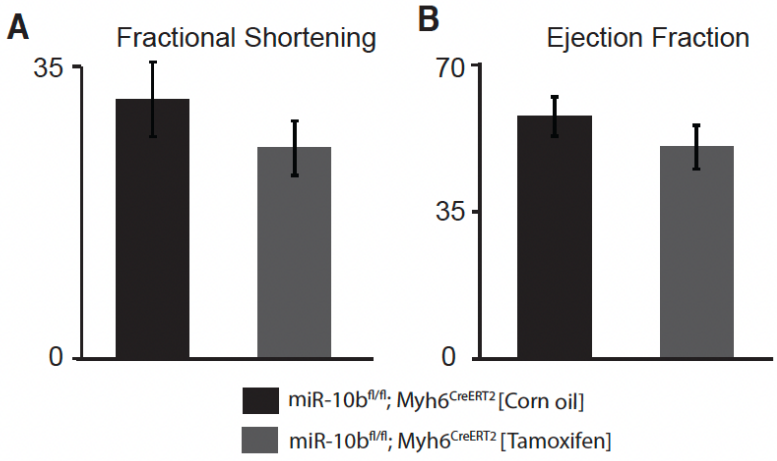


Figure 2.S3. No observable change in ejection fraction or fractional shortening after the removal of miR-10b.

Chapter 2.4 Discussion

Our work is aimed at elucidating the relationship between miRNAs and atrial fibrillation, and it highlights the importance of miRNAs in the maintenance of homeostatic cardiac rhythm. We have identified candidate miRNAs with changed expression after the removal of *Tbx5*, a mouse model for atrial fibrillation. The small RNAs are collected prior to the onset of atrial fibrillation, which suggests the changes in miRNA candidates may identify important regulatory mechanisms of cardiac rhythm.

A closer inspection of miRNA families reveals a trend where the majority of the miR-154 family members have increased expression after the removal of *Tbx5.* Interestingly, many of the family members are located in the intronic region of *Mirg* (miRNA containing gene) including miR-154, miR-369, miR-377, miR-409, and miR-496a. The remaining up-regulated miRNAs (miR-300, miR-382, miR-494) are located upstream of the *Mirg*. Clustered miRNAs can have similar expression changes, as their proximity makes it likely they will be transcribed together. This phenomenon has been studied in various contexts, and may be important when considering potential therapeutic treatments [183]–[185]. For example, mice subjected to myocardial infarction were not rescued with LNAs targeting miR-34 alone, but LNAs inhibiting the miR-34 family prevented cardiac remodeling and improved cardiac function [184]. Therefore, elucidating the role of the miR-154 cluster may provide meaningful insight into the pathophysiology of atrial fibrillation, as this cluster of miRNAs is up-regulated in our mouse model. Previous publications have also shown increased expression of miR-154 in two different mouse models, AngII-treatment and transverse aortic constriction (TAC) surgery [186], [187]. This suggests that the activation of miR-154 is not unique to atrial fibrillation, but rather it can be generalized to different cardiac insults and part of a shared regulatory network. In these studies the inhibition of miR-154 resulted

in an attenuation of cardiac hypertrophy and cardiac remodeling, implicating it as a driver of cardiac remodeling [186], [187]. The results for miR-154 are fascinating and make it a promising target for therapeutic treatments, but further elucidating the nearby cluster miRNAs may also provide meaningful insight into atrial fibrillation and potentially other cardiac diseases.

While the miR-154 family is a fascinating cluster of miRNAs, our initial efforts were spent on the miR-10b family. The miR-10 family has decreased expression after the removal of *Tbx5*, with miR-10b having the most down-regulation within the family. Our decision to evaluate this miRNA was influenced by the results of the high-throughput screen. The evaluation of our candidates in a high-throughput screen narrowed our list of candidates to those that may be important for atrial rhythm. Our miRNA sequencing identified 162 candidate miRNAs that have changed expression after the removal of *Tbx5*, and to evaluate each candidate in a traditional sense would have been an arduous task. The implementation of this high-throughput methodology was instrumental in efficiently evaluating many candidates, and also identifying candidates that may have an arrhythmogenic phenotype in an *in vitro* setting.

miR-10b is located within the Hox clusters, and is involved in various forms of cancer with increased expression in metastatic cancers [188], [189]. In the high-throughput screen the overexpression of miR-10b resulted in a prolonged action potential, and a high arrhythmia index score implies it plays an important role in cardiac rhythm. miR-10b has also been shown to have a role in cardiomyocytes, as miR-10a and miR-10b repress the expression of *Tbx5* by targeting the 3’ UTR [179]. In their studies, the repression of the *Tbx5* 3’UTR with miR-10a and miR-10b in a luciferase reporter assay is eliminated when the miR-10 binding sites are mutated; the regulation of *Tbx5* by miR-10a and miR-10b was confirmed in a western blot with miR-10b mimics and inhibitors [179]. Our laboratory also evaluated the ability of miR-10a/miR-10b mimics and

inhibitors to repress *Tbx5* expression in HL-1 cardiomyocytes, and confirmed this relationship with qPCR results (Figure 2.4C). The TBX5 ChIP experiment confirmed miR-10a and miR-10b are direct targets of TBX5, which corroborates the decreased expression of miR-10 after the removal of *Tbx5* in our mouse model (Figure 2.4D). This novel finding is fascinating as it demonstrates TBX5 drives the expression of these miRNAs, the miRNAs in turn regulate the expression of TBX5. This feedback mechanism helps fine tune TBX5 gene expression in order to maintain stable gene expression. This mechanism of regulation may provide insight into the phenomenon where the removal of exon 3 of *Tbx5* in our mouse model leads to increased transcription of the remaining exons. The absence of the third exon leads to a non-functional protein, but there is an increased level of the remnant transcript. We have evidence to support the notion that in the absence of a functional TBX5 protein, miR-10 also has decreased expression, which leads to less repression of the *Tbx5* transcript at the 3' UTR. Therefore, decreased expression of miR-10 leads to increased TBX5 expression, but in our mutant model it increased expression of a mutant, non-functional *Tbx5* transcript. Overall, our studies have corroborated the conclusion that miR-10a and miR-10b repress the expression of *Tbx5*, but more importantly provided evidence for a feedback mechanism that maintains stable levels of TBX5, which is necessary to maintain stable expression levels of the downstream ion channels and cardiac structure genes.

Cardiac rhythm is a dose sensitive process that requires fine tuning of gene expression to maintain homeostatic levels of proteins. We have previously shown that the absence of *Tbx5* leads to spontaneous, sustained atrial fibrillation, and a gain-of-function TBX5 mutant also results in paroxysmal atrial fibrillation [84], [118]. In each instance the expression levels of TBX5 are modulated to throw off the balance of the downstream target gene expression. Overexpression of miR-10b in the high-throughput *in vitro* experiment resulted in an arrhythmogenic phenotype, and

the *in vivo* miR-10b deficient mice were susceptible to atrial fibrillation. These observations suggest that modulating the expression of these regulatory networks in either direction can have detrimental effects on the maintenance of stable gene expression.

The high-throughput screen identified miR-10b as a candidate miRNA important in atrial rhythm, which we have now corroborated with *in vivo* and *in vitro* experiments. This is a testament to the ability of the high-throughput screen to identify functionally important candidates. This suggests that the other candidates identified in the screen deserve further investigation. For example, after the removal of *Tbx5* there is an increase in expression of miR-1231. The high-throughput screen determined overexpression of miR-1231 had a shortened action potential and a very high arrhythmia index calculation indicative of arrhythmogenic qualities in the *in vitro* study. The overexpression of miR-1231 in the high-throughput screen also showed ectopic beats when the action potentials were evaluated. miR-1231 has previously been shown to be up-regulated in ischemic hearts, and *Cacna2d2* is the direct target of this miRNA [190]. Knockdown of the miRNA resulted in an improvement of the electrophysiology defects seen in ischemic hearts. provides further proof that the high-throughput screen is an effective method to identify miRNAs involved in the regulation of cardiac rhythm [190]. miR-21 is another example of a miRNA that was up-regulated in our miRNA profiling experiment and identified as a promising candidate for involvement in cardiac rhythm maintenance. It is a very well-studied miRNA and has been evaluated in a variety of cardiac myopathies [191]. miR-1231 and miR-21 are two of the top three candidates in the results of the high throughput screen shown in Figure 2.S1. This provides evidence that the high-throughput screen is an effective method to identify miRNAs involved in the regulation of cardiac rhythm. With this in mind, the miRNA with the highest arrhythmia index score is miR-1193, which may warrant further investigation into its role in the heart and

maintaining atrial rhythm. This miRNA has been evaluated for its change in expression in different contexts of cancer, but in the heart its role is unexplored. The use of *in silico* prediction of miRNAs may provide preliminary knowledge regarding the target genes regulated by this miRNA. We hypothesize that miR-1193 is targeting genes important in cardiac rhythm, and modulation of its expression in the heart will provide an arrhythmogenic phenotype. Based on the our findings and the existing literature for the miRNA candidates we identified in this experiment with a significant Arrhythmia Index score, we posit these miRNAs are promising candidates to pursue in future studies evaluating miRNAs involved in cardiac rhythm.

In summary, as we continue to elucidate the role of miRNAs and their requirement for normal cardiac rhythm, it will further refine our understanding of the regulatory mechanisms required to maintain stable gene expression in the heart. Our small RNA profiling experiment identified differentially miRNAs in an atrial fibrillation model, and the evaluation of those candidates in a high-throughput screen narrowed the scope of those that may have an arrhythmogenic phenotype. Our studies validated a single candidate, miR-10b, but a literature search provided evidence for a role of several candidates in regulating genes important in cardiac rhythm. The relationship between TBX5 and miR-10b provides a feedback mechanism by which the miRNA can regulate the expression of *Tbx5* in a dose-dependent manner to function as a buffer to maintain stable gene expression. Future studies are warranted to identify the gene targets and understand the functional roles of the remaining candidates identified in the small RNA profiling experiment and high-throughput screen.

Chapter 2.5 Acknowledgments

The high-throughput experiments evaluating the arrhythmogenic phenotypes with miRNA overexpression were completed by Michael Yu from the laboratory of Alexandre Colas (Development, Aging and Regeneration Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037). The small RNA-seq, electrophysiology studies, and ChIP experiments were performed in a collaborative effort by Rangarajan D. Nadadur, Margaret Gadek, Kaitlyn M. Shen, Carlos Perez-Cervantes, Brigitte Laforest, and Christopher R. Weber.

CHAPTER 3: SHARED REGULATORY NETWORK IN ATRIAL FIBRILLATION AND HEART FAILURE

Chapter 3.1 Introduction

There is a strong epidemiologic link between heart failure and atrial fibrillation, as up to half of all heart failure patients ultimately acquire atrial fibrillation. These two cardiac diseases are associated with increased incidence of each other, and in conjunction there is an increase of morbidity and mortality compared to each disease alone. Atrial fibrillation and heart failure share common risk factors, including age, high blood pressure, diabetes, and obesity. Both cardiac diseases also share structural and physiologic endpoints, including atrial remodeling. Therefore, examining the molecular mechanisms driving the pathophysiology is crucial to understanding the mechanistic basis of their shared incidence. Fortunately, we have mouse models that allow us to explore the pathophysiology of these cardiovascular diseases. Our lab has previously shown that the deletion of TBX5 results in spontaneous, sustained atrial fibrillation, demonstrating the instrumental role of *Tbx5* in the maintenance of cardiac rhythm. A pressure overload-induced heart failure mouse model is accomplished through a transverse aortic constriction (TAC) surgery. Comparison of these two mouse models provides valuable insight into their shared pathophysiology.

A proper functioning heart has a stable, rhythmic heartbeat that is regulated by the expression of a system of ion channels, gap junction proteins, ions, and many other critical components. When there is a disruption in this coordinated electrical activity, it can lead to an irregular heartbeat, or a cardiac arrhythmia. Atrial fibrillation is the most common type of cardiac arrhythmia that is characterized by an irregular heartbeat due to chaotic electrical signals in the

atria, and unfortunately it is a growing epidemic worldwide. It is estimated that atrial fibrillation affects around 2.5-5.1 million people in the United States, but it is projected to continue increasing [11], [12]. This is not an epidemic that is unique to the Unites States, as global estimates for atrial fibrillation are ~33.5 million individuals [13]. Due to the considerable number of individuals that are affected by this medical condition, it is not surprising that an estimated $26 billion dollars are spent annually in the United States to treat atrial fibrillation [192]. The current atrial fibrillation treatments include beta blockers, ion channel blockers, catheter ablations, and pacemakers. These therapeutic treatments are used to manage the symptoms and prevent other cardiac complications, but can be associated with significant side effects and risks [14]. Therefore, we believe understanding the molecular mechanisms driving the pathophysiology of atrial fibrillation may provide more effective therapeutic strategies to manage and treat the disease. Identifying the most effective therapeutic treatment is crucial for patients with atrial fibrillation as they have an increased risk of stroke, heart failure, and death. Fortunately, we are able to replicate aspects of atrial fibrillation in mouse models, which allows us to explore gene regulatory mechanisms and signaling pathways that are changed in disease [193]. For our studies we compared *Tbx5*-deletion (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) and control (*R26*^{CreERT2}) mice, a mouse model for atrial fibrillation previously published by our laboratory [64], [118]. The adult-specific deletion of *Tbx5* in the mouse leads to an absent p-wave, irregular heartbeat, and prolonged action potentials, all canonical markers of atrial fibrillation. Transcriptional profiling studies in these mice revealed TBX5 drives the expression of a gene regulatory network important in cardiac rhythm, including the regulation of calcium handing genes.

Heart failure is a growing epidemic that is characterized by the inability of the heart to pump a sufficient amount of blood to meet the metabolic needs of the body. Heart failure is

estimated to affect around 6 million people in the United States, and its prevalence is projected to increase to more than 8 million people by 2030 [121], [127]. This phenomenon is not unique to the Unites States, as heart failure is estimated to affect around 23-26 million people worldwide [125], [126]. As the incidence of heart failure is increasing, the cost to treat heart failure is estimated to also increase from $20.9 billion annually to $53.1 billion by 2030 [121], [127]. There are several different stages of heart failure, so treatments are dependent on the progression of the disease and can include beta blockers, angiotensin II receptor blockers, and surgery. Fortunately we are able to simulate this cardiac disease in animal models, in order to elucidate the molecular mechanisms progressing the cardiac disease. For our studies our collaborators from the lab of David Park performed the transverse aortic constriction (TAC) surgery, a pressure overload-induced cardiac hypertrophy and heart failure mouse model [194]. Following TAC surgery there is a compensated cardiac hypertrophy response as a mechanism to deal with the increase in pressure, and it will eventually lead to heart failure due to the maladaptive changes in the heart. This mouse model has been extensively used in the field to explore the pathophysiology of heart failure, and to identify critical regulatory during the progression of disease.

While there is a strong epidemiologic link between heart failure and atrial fibrillation, the relationship between these two cardiac diseases has only been partially elucidated. It is estimated that 40% of individuals with atrial fibrillation or heart failure will also be diagnosed with the reciprocal condition [121], [122]. Therefore, atrial fibrillation can lead to heart failure, and heart failure can lead to atrial fibrillation. The previous paradigm suggested heart failure has increased fibrosis from structural remodeling to disrupt the electrical connectivity in the heart leading to atrial fibrillation [195]. The irregular heartbeat in atrial fibrillation results in ineffective systole and insufficient amounts of blood are pumped, which can lead to heart failure. A diagnosis of atrial

fibrillation or heart failure has an increased risk to mortality, but the prognosis is much worse for those that are diagnosed with both. Their comorbidity has only been examined superficially, and their coexistence is predominately been attributed to the many shared risk factors associated with both cardiac diseases. Age, coronary artery disease, myocardial infarction, high blood pressure, and obesity are all risk factors for both atrial fibrillation and heart failure. While these cardiac disease are influenced by these risk factors, we believe there is a shared gene regulatory network that can be attributed to the comorbidity of these two diseases. Atrial fibrillation and heart failure both lead to cardiac remodeling processes with fibrosis, inflammation, and electrophysiological changes. To understand the shared pathophysiology of atrial fibrillation and heart failure, we compared the gene regulatory networks of these two cardiac perturbations.

The comparison of the transcriptional changes in the *Tbx5*-deletion mouse model and the Transverse Aortic Constriction (TAC) mouse model revealed remarkable correlation between the differentially expressed genes in these disparate disease models. Our analysis identifies candidate transcription factors that may be driving the pathophysiology of these cardiac diseases. The non-coding regions of the genome also revealed a positive correlation in the comparison of *Tbx5*-deletion and TAC mice. It is well understood that active enhancers produce noncoding RNAs (ncRNAs) transcripts, therefore our analysis reveals a shared regulatory network. We focused our efforts on a down-regulated enhancer-borne ncRNA in the *Klf15* locus, and the associated open chromatin peak identified an active enhancer in cardiomyocytes that is directly regulated by TBX5. *Klf15* is reported to be a repressor of cardiac hypertrophy, therefore we provide evidence that TBX5 has an additional role in cardiac homeostasis. We also identified a disease-specific ncRNA in the *Sox9* locus that is up-regulated in cardiac fibroblasts during disease conditions. Overall, the comparisons of the coding and the non-coding regions of the atrial fibrillation and heart failure

mouse models identified a shared gene regulatory network driving the pathophysiology of these cardiac diseases.

Chapter 3.2 Material and Methods

Mouse model for atrial fibrillation

The mice used in these experiments were raised in accordance with the Guide for the Care and Use of Laboratory Animals. Generation of the conditional *Tbx5* deletion was completed with a floxed allele, and was previously described by our laboratory and other [64], [118], [159]. The third exon of *Tbx5* was excised with a tamoxifen inducible Cre recombinase driven by the R26 locus. $Tbx5^{fl/fl}$;$R26^{CreERT2}$ and $R26^{CreERT2}$ mice were subjected to a tamoxifen regiment over 3 days at 6-10 weeks of age as previously described [118]. All experiments were approved by the University of Chicago Institutional Animal Care and Use Committee (IACUC).

Mouse model for pressure overload by transverse aortic constriction (TAC) procedure

The transverse aortic constriction (TAC) surgery was performed for a pressure overload-induced heart failure mouse model. The TAC mice were generated by our collaborators in the laboratory of David Park at New York University (NYU). Constriction of the transverse thoracic aorta was performed with a silk suture tightened against a 27-gauge needle. Sham mice underwent the same surgery but without TAC.

Control and Tbx5-mutant coding RNA-seq library preparation and sequencing

Left atrial tissue from $R26^{CreERT2}$ (control) and $Tbx5^{fl/fl}$;$R26^{CreERT2}$ (*Tbx5*-mutant) mice was dissected one week after tamoxifen regiment, and mechanically homogenized in TRIzol Reagent. Total RNA was prepared with a TRIzol-based extraction and isolated using a Qiagen RNeasy Mini Column as previously described [118], [159]. Library preparation and sequencing was performed at the University of Chicago Genomics Core Facility. The RNA-seq dataset can be downloaded on the GEO database (GSE129503) [118].

Sham and TAC coding RNA-seq library preparation and sequencing

Left atrial tissue from sham and TAC mice were removed, prepared with a TRIzol-based extraction protocol, and isolated using a Qiagen RNeasy Mini Column. PolyA selection was performed using the Lexogen CORALL PolyA selection module, and subsequently library preparation was completed with the Lexogen CORALL Total RNA Library Kit. The libraries were sequenced on the Illumina NovaSeq flowcell. Samples were sequenced 50bp paired-ended at 12-22 million reads per replicate. The reads were mapped to mm10 using STAR v2.5.3 (UCSC mm10 annotation). Genomic read alignments (BAM files) were filtered by MAPQ score greater than 30 with samtools v1.9, and counts were generated for each sample with HTSeq-counts (v.0.11.2) with mm10 known gene annotation from UCSC [196]. Raw counts were normalized by Relative Log Expression (RLE) in DESeq2, and genes with counts larger than 10 in at least 3 samples, which is the sample size of one condition, are kept for downstream analysis. Then, raw counts of the expressed genes are used by the default DESeq2 analysis pipeline to identify differentially expressed genes between sham and TAC [197]. Genes with a |log2 fold change| greater than 0.5, and adjusted p-value less than 0.05 are considered significantly differentially expressed.

Control and Tbx5-mutant non-coding RNA-seq library preparation and sequencing

Left atrial tissue was removed from *R26*^{CreERT2} (control) and *Tbx5*$^{fl/fl}$;*R26*^{CreERT2} (*Tbx5*-mutant) mice and total RNA was extracted using a TRIzol protocol as previously described by our laboratory [159]. RiboZero purification and oligo-dT depletion was performed to enrich for non-coding transcripts. RNA libraries were generated using Illumina's TruSeq RNA Library Prep Kit v2. Samples were sequenced with 50bp single-ended reads on the HiSEQ4000 instrument, and the samples were sequenced to a depth of 65-95M reads per samples [159]. The reads were mapped to mm10 using STAR v2.5.3. Genomic read alignments (BAM files) were filtered by MAPQ score greater than 30 with samtools v1.9. Transcriptome-guided *de novo* assembly was performed using

stringtie/1.3.3. The data has been previously published by our laboratory in eLife by Xinan H. Yang and Rangarajan D. Nadadur [159].

Sham and TAC non-coding RNA-seq library preparation and sequencing

The left atria was removed from sham and TAC mice, and total RNA was isolated using a TRIzol extraction protocol as previously described by our laboratory [118], [159]. After oligo-dT depletion and RiboZero purification, RNA libraries were prepared according to Lexogen's instructions for CORALL Total RNA Library Kit. The polyA-depleted fractions were sequenced with 50bp paired-end to a depth of 88-99M reads per replicate. To appropriately compare with the *Tbx5*-mutant samples, only the first reads were mapped to mm10 using STAR v2.5.3. Genomic read alignments (BAM files) were merged and filtered by MAPQ score greater than 30 with samtools v1.9. Transcriptome-guided *de novo* assembly was performed using stringtie/1.3.3.

Comparison of non-coding RNA transcripts in Tbx5-mutant and TAC

Stringtie merge was performed across all samples in *Tbx5*-mutant and TAC experiments to obtain a comparable assembly. Transcripts that overlap with protein-coding genes defined by GENCODE v25, transcripts with a 'NM' annotation, or isoform transcripts with a 'XM' label were removed from the assembly. Runoff transcripts, defined as 5k bp downstream of the 3' end of a protein coding gene, were also removed from the assembly. Counts for each sample are generated with HTSeq-counts (V.0.11.2) [196]. To remove low expressed transcripts, transcripts with CPM greater than 10 or 11 for *Tbx5*-mutant experiment and TAC experiment, respectively, in at least 3 samples are kept in downstream analysis. The raw counts of these expressed transcripts are analyzed in edgeR and identify differentially expressed transcripts between the control and experimental condition.

Association of differentially expressed genes with ncRNA-enhancer candidates

A ncRNA candidate within 1M bp from a differentially expressed gene are considered to be a potential regulator of that gene of interest.

Generation of ATAC-seq dataset and analysis

Left atria and cardiomyocyte ATAC-seq (Assay for Transposase-Accessible Chromatin sequencing) datasets were generated using Isolation of Nuclei Tagged in specific Cell Types (INTACT). The Cre-loxP system is used to express a GFP-tagged SUN1 nuclear membrane protein allowing the purification of cell-type specific nuclei. To generate the left atrial ATAC-seq dataset we used the $R26^{\text{CreERT2}}$;$R26^{Sun1\text{-GFP}}$ mouse line to isolate the nuclei of all the cell types in the left atria, as R26 locus is ubiquitously expressed. To generate the cardiomyocyte ATAC-seq dataset we used the $Myh6^{\text{CreERT2}}$; $R26^{Sun1\text{-GFP}}$ mouse line as *Myh6* is a cardiomyocyte-specific gene. The purified nuclei were used for the transposition reaction. The fibroblast-specific ATAC-seq dataset was generated using a published protocol to isolate primary adult mouse cardiac fibroblasts based on differential adhesive properties of the fibroblasts [198]. The isolated cardiac fibroblasts were resuspended in a cold lysis buffer (10mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL CA-630), and we subsequently performed the transposition reaction. The transposition protocol was performed based on the published ATAC-seq protocol. Libraries were generated in accordance with the Illumina Nextera DNA Library Prep Kit protocol. The left atria and cardiomyocyte ATAC-seq datasets were sequenced with 50bp paired-end reads and the fibroblast ATAC-seq dataset was sequenced with 50bp single-end reads. The samples were sequenced on the Illumina HiSeq machine at the University of Chicago Genomics Core Facility. ATAC-seq reads were mapped to mm10 with bowtie2/2.3.2. Genomic read alignments (bam files) were filtered by MAPQ score larger than 30 to ensure a certain level of unique mapping with samtools v1.5, and then sorted. Peaks were called using ATAC-seq mode of Genrich/0.5 removing

PCR duplicates and filtering with p-value 0.05 for left atrial and cardiomyocyte ATAC-seq, and 0.02 for fibroblast ATAC-seq.

Gene ontology enrichment analysis

Gene ontology analysis was performed using Metascape (https://metascape.org) utilizing both the single gene list and multiple gene list functions.

Statistical analysis of the significant overlap between the differentially expressed genes and transcription factors of two mouse models

Fisher's exact tests were used to calculate statistical significance and the odds ratios of overlaps between the differentially expressed genes and the differentially expressed transcription factors between the *Tbx5*-mutant and TAC experiments. The null hypothesis is defined as a specific group of genes are independent in the two experiments, and a p-value smaller than 0.05 is considered significant to reject this null hypothesis suggesting there is a non-random association between the two experiments.

Ingenuity Pathway Analysis (IPA)

The DESeq2 result tables were used as input for Ingenuity Pathway Analysis (IPA) for both the *Tbx5*-deletion and TAC experiments ($|log2FoldChange|>0.5$, $p.adj<0.05$). Expression analysis was performed using log2 fold change to calculate the z-score, which is subsequently used to compare with human and mouse experimentally observed pathways. Heatmap plot is sorted by z-score of canonical pathways with a cutoff of 3 and -1 for predicted activated and repressed pathways, respectively.

Motif enrichment analysis

Known and de novo motif scanning was performed with HOMER. Candidate enhancers identified as ATAC-seq peaks that overlap (1k bp upstream and 50 bp downstream) the 5' end of candidate non-coding RNAs.

TBX5 ChIP-seq library

Chromatin extract was prepared from adult mouse cardiac tissue. The tissue was cross-linked in PBS containing 1% formaldehyde for 10 minutes at room temperature, and subsequently quenched by 12 mM glycine. The cross-linked samples were incubated in Lysis Buffer 1 ((50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100) and proteinase inhibitors on ice for 10 min. Lysis Buffer 3 (1 mM EDTA, 0.5 mM EGTA, 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1% sodium deoxycholate, 0.5% N-lauroyl sarcosine) was subsequently added to the samples, and sonicated for 15 minutes. The chromatin extract was incubated with anti-TBX5 antibody (sc-17866, Santa Cruz Biotechnology) overnight at 4°C. The samples were incubated with Dynabeads to collect the immunocomplex, and rinsed with buffers. Elution buffer (1% SDS in TE with 250 mM NaCl) is added to the immunocomplexes to elute the captured chromatin at 65°C. DNA was purified after the completion of RNase treatment, proteinase K treatment, and reverse cross-linking. Sequencing libraries were generated from ChIP and input DNA using NEBNext Ultra DNA Library Prep Kit (New England Biolabs). The samples were sequenced by the University of Chicago Genomics Core Facility using Illumina HiSeq instruments.

Detecting chromatin loops from micro-C dataset

Library preparation and sequencing was completed by our collaborators in the Alex Ruthenburg laboratory at the University of Chicago. Significant contacts were obtained from

Mustache v.1.2.0 using the contact matrix bin size of 1 kb and subsequently filtered for loops having at least a false discovery rate of 0.1 [199].

Isolation of cardiac primary fibroblasts and TGF-β treatment

Primary cardiac fibroblasts were isolated from the adult mouse heart based on differential adhesive properties of fibroblasts, as described in the published protocol [198]. Five adult mouse hearts were minced and digestion buffer was added to the tissue. Following digestion of the tissue, the cell suspension was plated into culture dishes and incubated at 37°C for 2 hours. The fibroblasts adhered to the culture dish and the supernatant was discarded. The adhered fibroblasts were further cultured for generation of RNA-seq, ncRNA-seq, and ATAC-seq genomic datasets. To evaluate the activation of fibroblasts in an *in vitro* context the cells were treated with TGF-β, a canonical stimulus for fibroblast activation. The cultured fibroblasts were treated with vehicle (control) or TGF-β (10ng/mL) and collected after 48 hours.

RNA-seq: Control and treated cardiac fibroblasts were prepared with a TRIzol-based extraction protocol, and isolated using a Qiagen RNeasy Mini Column. PolyA selection was performed using the Lexogen CORALL PolyA selection module, and subsequently library preparation was completed with the Lexogen CORALL Total RNA Library Kit. The libraries were sequenced on the Illumina NovaSeq flowcell. Samples were sequenced 50bp paired-ended at 16-20 million reads per replicate. The reads were mapped to mm10 using STAR v2.5.3(UCSC mm10 annotation). Genomic read alignments (BAM files) were filtered by MAPQ score greater than 30 with samtools v1.9, and counts were generated for each sample with HTSeq-counts (v.0.11.2) with mm10 known gene annotation from UCSC [196]. Raw counts were normalized by Relative Log Expression (RLE) in DESeq2, and genes with counts larger than 10 in at least 3 samples, which is the sample size of one condition, are kept for downstream analysis. Then, raw counts of the

expressed genes are used by the default DESeq2 analysis pipeline to identify differentially expressed genes between sham and TAC [197]. Genes with a |log2 fold change| greater than 0.5, and adjusted p-value less than 0.05 are considered significantly differentially expressed. The samples were sequenced on the Illumina Nova-Seq machine at the University of Chicago Genomics Core Facility.

ncRNA-seq: The total RNA was extracted from control and TGF-β treated fibroblasts using a TRIzol extraction protocol as previously described by our laboratory. After polyA selection and RiboZero purification, the remaining RNA was prepared according to Lexogen's instructions for CORALL Total RNA Library Kit. This polyA-depleted fraction was sequenced with 50bp paired-end to a depth of 14-18M and 50bp single-end to a depth of 34-82M reads per replicate. The samples were mapped to mm10 using STAR v2.5.3 separately. Genomic read alignments (BAM files) were merged and filtered by MAPQ score greater than 30 with samtools v1.9. Transcriptome-guided *de novo* assembly was performed using stringtie/1.3.3. The samples were sequenced on the Illumina Nova-Seq machine at the University of Chicago Genomics Core Facility.

ATAC-seq: The isolated cardiac fibroblasts were resuspended in a cold lysis buffer (10mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL CA-630), and subsequently performed the transposition reaction. The transposition protocol was performed based on the published ATAC-seq protocol. Libraries were generated in accordance with the Illumina Nextera DNA Library Prep Kit protocol. The samples were sequenced on the Illumina Nova-Seq machine at the University of Chicago Genomics Core Facility. 50bp paired-end ATAC-seq reads were mapped to mm10 with bowtie2/2.3.2. Genomic read alignments (bam files) were filtered by MAPQ score larger than 30 to ensure a certain level of unique mapping with samtools v1.5, and

then sorted. Peaks were called using ATAC-seq mode of Genrich/0.5 removing PCR duplicates and filtering with p-value 0.05.

Luciferase Assay

Candidate enhancers were amplified from mouse genomic DNA, and cloned into pGL4.23 luciferase expression vector with a minimal promoter. The sequence was confirmed by sequencing. Mutagenesis of the transcription factor motifs was made from gBlocks DNA fragments (IDT), and cloned into luciferase expression vectors. HL-1 cardiomyocytes and 3T3 fibroblasts were co-transfected with the luciferase expression vector and pRL control (renilla) vector using Lipofectamine 3000. The cells were lysed and assayed after 48 hour treatment using the Dual-Luciferase Reporter Assay Kit (Promega). Transfection protocols have previously been described by our laboratory [159].

Chapter 3.3 Results

Shared transcriptional response in atrial fibrillation and heart failure mouse models

To understand the shared molecular mechanisms between atrial fibrillation and heart failure, we performed RNA profiling in mouse models for these cardiac perturbations. Our lab has previously demonstrated that removal of *Tbx5* leads to atrial fibrillation, and the transcriptional profiling data from this publication was used for our comparative analysis [118], [159]. The left atria from *Tbx5*$^{fl/fl}$; *R26*^{CreERT2} and *R26*^{CreERT2} mice were collected 1-week after tamoxifen treatment, which is prior to the onset of atrial fibrillation. Differential gene expression analysis between the two genotypes revealed 1,914 down-regulated genes, and 1,913 up-regulated genes ($|log2FoldChange|>0.5$, $p.adj<0.05$, Figure 3.1 A-C). Gene ontology analysis was completed for the down-regulated and up-regulated genes separately to explore the variety of genes that were repressed and activated. The gene ontology analysis of significantly down-regulated genes after the removal of *Tbx5* are related to heart contraction, which is expected due to the fact that TBX5 directly regulates the expression of ion channels (Figure 3.1A) [118]. Interestingly, the gene ontology analysis for the up-regulated genes reveals an enrichment for terms related to disease processes, including inflammatory response, leukocyte migration, and extracellular matrix organization (Figure 3.1C). While a previous publication from our lab provided insight into the role of *Tbx5* in directly regulating cardiac rhythm, a more in-depth analysis of the genes that are upregulated in response to *Tbx5* deletion suggests the involvement of other cell types, including immune cells and fibroblasts.

We also completed transcriptional profiling for the left atria in a control sham surgery compared to transverse aortic constriction (TAC) pressure overload mouse model (Figure 3.1D-F). Two weeks after the surgeries, the tissues were collected and processed for RNA-seq.

Differential gene expression analysis between the sham and TAC identified 1,926 down-regulated genes and 1,931 up-regulated genes (|log2FoldChange|>0.5, p.adj<0.05, Figure 3.1B). The gene ontology analysis for the up-regulated genes in TAC were similar to those seen in the TBX5 KO RNA-seq, and included enrichment for disease terms including "extracellular matrix organization" and "hemostasis". The gene ontology analysis for the down-regulated genes in TAC banding revealed terms related to metabolic processes including "TCA cycle and respiratory electron transport" and "mitochondrion organization". The enrichment of these terms is expected in a TAC RNA-seq, as they have been previously implicated in other transcriptional profiling experiments in the heart [135], [200], [201].

While the we identified similarities in the up-regulated gene ontology terms in both atrial fibrillation and heart failure, we wanted to directly compare the transcriptional changes between these two cardiac perturbations. A comparison of differential gene expression for *Tbx5* KO (x-axis) and TAC (y-axis) reveals a positive correlation of 0.82, thus highlighting a shared transcriptional response (Figure 3.1G). A direct overlap of the significant differentially expressed genes reveals 770 shared down-regulated genes (OR=7.5; p<2.2e-16; Figure 3.1H, left) and 1,242 shared up-regulated genes (OR=23.6; p<2.2e-16; Figure 3.1H, right). These overlapped genes reveal a network of genes that are shared in the two cardiac diseases. While we focus on identifying the shared regulatory mechanisms of atrial fibrillation and heart failure, there are unique genes for each of the cardiac disease mouse models that can be examined in future investigations.

While gene ontology analysis for each individual dataset was crucial in identifying significant biological processes and pathways, it is also instrumental to do comparative gene ontology analysis in order to identify similarities and differences between the two datasets (Figure 3.1I-J). This analysis of the down-regulated genes in the TAC and *Tbx5* KO mouse models reveals

a shared enrichment for terms related to cardiac conduction, including “Muscle contraction (R-MMU-397014)” and “regulation of transmembrane transport (GO:0034762)” (Figure 3.1I). While our previous research publication has demonstrated the importance of *Tbx5* in regulating the expression of ion channels, we have now confirmed TAC mice also have down-regulation of genes involved in cardiac conduction. Therefore, both atrial fibrillation and heart failure lead to electrical remodeling in the atria, highlighting electrical remodeling as a shared disease process. The down-regulated genes in the TAC mouse model are also enriched for terms related to metabolic processes, including “The citric acid (TCA) cycle and respiratory electron transport (R-MMU-1428517)” (Figure 3.1I). This enrichment is specific to the TAC mouse model and absent in the *Tbx5* KO mouse model as shown in the heatmap. It is well established that heart failure leads to profound metabolic abnormalities, so its predominance in the TAC mouse model is expected [202], [203]. When we expand this to the top 100 terms, we can continue to see the same pattern of shared enrichment in conduction terms, and metabolic terms that are predominantly in the TAC mouse model (Figure 3.S1A). The discrepancy of metabolic terms compared to the *Tbx5* KO mouse model highlights some of the differences between these two cardiac diseases.

When examining the gene ontology analysis for the up-regulated genes in the TAC and *Tbx5* KO mouse models, it is apparent that all the enriched terms are shared between the two cardiac diseases, including “extracellular matrix organization (GO:0030198)”, “regulation of cytoskeleton organization (GO:0051493)”, and “supramolecular fiber organization (GO:0097435)” (Figure 3.1J). These enriched terms are most likely indicative of the transcriptional changes that occur during fibrosis, an important process during pathological remodeling characterized by the fibroblast activation with increased amount of extracellular matrix proteins being expressed. While it is apparent there are distinctions in the down-regulated genes

when comparing atrial fibrillation and heart failure, it is fascinating that all the enriched terms for the up-regulated genes are shared among the two cardiac perturbations. When we expand this to the top 100 terms enriched in these two datasets, we continue to see a completely shared enrichment of disease terms for both models of cardiac disease (Figure 3.S1B). The enriched terms are also inclusive of terms related to immune processes including "myeloid leukocyte activation (GO:0002274)", "regulation of alpha-beta T cell activation (GO:0046632)", and "regulation of immune effector process (GO:0002697)" (Figure 3.S1B). The shared immune gene ontology terms suggests there is an immune infiltration response during these cardiac disease processes.

While there are distinct regulatory mechanisms in the comparative gene ontology for the down-regulated genes, the shared regulatory mechanisms in the differentially expressed genes suggest a shared pathophysiologic response in atrial fibrillation and heart failure. To evaluate this further, we performed comparative Ingenuity Pathway Analysis (IPA) to identify putative shared pathways in cardiac disease (Figure 3.S2). The pathways in the down-regulated genes include Calcium Signaling, Apelin Muscle Signaling Pathway, and FGF Signaling; the pathways in the up-regulated genes include Hepatic Fibrosis Signaling Pathway, Wound Healing Signaling Pathway, and Leukocyte Extravasation Signaling. Gene ontology and pathway analysis have made it clear that there is a conservation in the regulatory network controlling cardiac disease, but future investigations will allow us to further examine specific pathways of interest.

In order to identify the transcriptional regulators driving the pathophysiology in cardiac disease, we overlapped the differentially expressed transcription factors in the *Tbx5* KO and TAC mouse models (Figure 3.1K). Transcription factors are DNA-binding proteins that regulate gene expression, therefore this comparison identified 142 candidate transcription factors that could be responsible for changing the expression of genes in a cardiac disease context. Interestingly, closer

examination of the 142 differentially expressed genes in *Tbx5* KO and TAC reveals that 135 of the transcription factors share directionality in their misexpression as shown in the heatmap (Figure 3.1L). Several of the candidate transcription factors have already been implicated in cardiac disease, including *Meox1* and *Runx1*. The shared transcriptional regulators implicates a shared mechanism of disease in atrial fibrillation and heart failure, and also identifies putative candidates that are responsible for driving the shared regulatory mechanisms in cardiac disease.

Positive correlation in differentially expressed ncRNAs implicates a shared regulatory network in atrial fibrillation and heart failure.

While we have shown that atrial fibrillation and heart failure mouse models have shared transcriptional changes, we wanted to evaluate whether these two cardiac diseases have shared enhancers in their regulatory networks. Non-coding regions of the genome are pervasively transcribed, particularly at enhancer regions. The differential deep sequencing of ncRNA transcripts can be used to identify highly active enhancers in a context dependent manner. We utilized this approach in both the *Tbx5* KO atrial fibrillation and TAC pressure overload mouse models. Our laboratory has previously published the non-coding transcriptional profiling of the polyA-depleted RNA transcripts in the control ($R26^{CreERT2}$) and *Tbx5* KO ($Tbx5^{fl/fl}$; $R26^{CreERT2}$) left atria [159]. This analysis identified enhancer-borne ncRNAs at ion channels, including *Atp2a2* and *Ryr2*, genes crucial for maintaining calcium homeostasis. In collaboration with the Dr. David Park laboratory at NYU, we also performed non-coding profiling in the left atria of sham and TAC mice. Dr. Park and his colleagues completed the surgeries and collected the tissue two weeks post-surgery, and our laboratory generated the libraries for sequencing.

De novo transcript assembly identified around 20,000 non-coding transcripts in the *Tbx5* experiment, and of these transcripts 3,895 transcripts were differentially expressed in the absence

of *Tbx5* (Figure 3.2A (center); |log2FoldChange|>0.5, p.adj.<0.1). The loss of *Tbx5* led to 2,013 down-regulated (log2FoldChange<0.5, p.adj.<0.1) ncRNA transcripts and 1,882 up-regulated (log2FoldChange>0.5, p.adj.<0.1) ncRNA transcripts. Gene ontology analysis was performed on the genes that are differentially expressed within a 1 MB distance from the ncRNA transcript. The down-regulated ncRNAs are near DE genes enriched for "regulation of striated muscle contraction (GO:0006942)" and "regulation of transmembrane transport (GO:0034762)" (Figure 3.2A (left)), while the up-regulated ncRNAs are near genes enriched for "Extracellular matrix organization (R-MMU-1474244)" and "Hemostasis (R-MMU-109582)" (Figure 3.2A (right)). The TAC surgery lead to 7,475 differentially expressed ncRNA transcripts (Figure3.2B (center); |log2FoldChange|>0.5, p.adj.<0.1). This pressure overload model leads to 4,333 down-regulated (log2FoldChange<0.5, p.adj.<0.1) ncRNA transcripts and 3,142 up-regulated (log2FoldChange>0.5, p.adj.<0.1) ncRNA transcripts. Gene ontology analysis was performed on the differentially expressed genes that are within 1 MB of the down-regulated ncRNAs, and it was enriched for "TCA cycle and respiratory electron transport (R-MMU-1428517)" and "mitochondrial gene expression (GO:0140053)" (Figure 3.2B (left)). The up-regulated ncRNAs were near genes enriched for gene ontology terms related to disease including "Extracellular matrix organization (R-MMU-1474244)" and "regulation of inflammatory response (GO:0050727)" (Figure 3.2B (right)). The gene ontology analysis suggests that the up-regulated ncRNA transcripts are near DE genes related to the cardiac disease processes in both atrial fibrillation and heart failure, while the down-regulated ncRNA transcripts are near DE genes related to conduction and metabolic processes.

While our analysis of the differentially expressed coding genes in the TAC and *Tbx5* KO mouse models reveals a shared gene regulatory network, we wanted to further evaluate whether

these two cardiac perturbations use the same regulatory elements to drive those gene expression changes. Previous research has shown that non-coding transcripts can be utilized as a tool to identify enhancer activity, we therefore compared the non-coding transcripts in the two cardiac disease mouse models. For a global analysis of the differentially expressed ncRNAs, we computed the correlation of the transcripts between the *Tbx5* KO and TAC ncRNA-seq (Figure 3.2D). Differential ncRNAs transcripts are highly correlated between the two cardiac disease models with a correlation of 0.82. The positive correlation highlights the shared transcription at the non-coding regions within the atrial fibrillation and pressure overload mouse models, and suggests there is a shared non-coding regulatory network that is activated in the stress of cardiac disease.

We wanted to take a closer look at the genomic locations of these shared non-coding RNA transcripts. Our laboratory has previously identified an enhancer-borne ncRNA, RACER, located upstream of Ryanodine Receptor 2 (*Ryr2*). Their work determined the enhancer was activated by TBX5, and the associated ncRNA, RACER, was chromatin enriched and functionally required for *Ryr2* expression. This non-coding RNA transcript was also identified as one of the shared down-regulated ncRNAs in atrial fibrillation and pressure overload mouse model. Examination of the *Ryr2* genomic locus shows the ncRNA present in the $R26^{\mathrm{CreERT2}}$ mice and the sham surgery mice, but in the $Tbx5^{fl/fl}$;$R26^{CreERT2}$ mice and TAC mice the transcript is absent (Figure 3.2F). This analysis has shown that this functional ncRNA is changed in both cardiac disease models, and also provides an example of a regulatory element that is shared in the atrial fibrillation and pressure overload mouse model.

Our previous work on the differentially expressed ncRNAs was focused on the transcripts that are directly regulated by TBX5 and are down-regulated following *Tbx5* removal, but in our new analysis we wanted to focus on the up-regulated ncRNA transcripts as well. We wanted to

identify ncRNAs that are activated in cardiac disease, with particular interest in those that are activated in both atrial fibrillation and pressure overload. Shown in the genomic tracks is a candidate ncRNA, *Wisper*, located upstream of *Wisp2* and has previously been published by the Ounzain and Pedrazzini laboratories [168]. This ncRNA is shown to regulate cardiac fibrosis after myocardial infarction, and its depletion *in vivo* attenuates fibrosis and improves cardiac function [168]. When we examine this locus in our mouse models, we demonstrate the transcript is present in the *Tbx5*$^{fl/fl}$;*R26*^{CreERT2} mice and TAC mice, but absence in the *R26*^{CreERT2} mice and the sham mice (Figure 3.2E). This highlights the ncRNA is not only active in myocardial infarction, but also pressure overload and atrial fibrillation, therefore the activation of this transcript occurs in various cardiac stresses and further validates there is a shared disease response.

It is well annotated that non-coding RNA transcripts produced from regulatory elements can be a reliable metric of enhancer activity [158]–[160], [162]–[164], [204]. Our analysis has identified differentially expressed ncRNA in atrial fibrillation and pressure overload, and we subsequently wanted to identify candidate enhancers. A hallmark characteristic of enhancer activity is the presence of open chromatin peaks, so we generated ATAC-seq datasets for two cell types in the heart. We utilized the INTACT (isolation nuclei tagged in specific cell types) method to isolate nuclei from specific cell populations based on the Cre-recombinase expression [205]. The isolated nuclei from a *R26*^{CreERT2} recombinase was used to generate a left atria ATAC-seq dataset, as Rosa26 locus is used for constitutive, ubiquitous expression in the mouse. We also generated a cardiomyocyte ATAC-seq dataset with a *Myh6*^{CreERT2} recombinase, as *Myh6* is a cardiomyocyte-specific gene that encodes the heavy chain subunit of cardiac myosin in cardiomyocytes. Examination of the *Myh6*/*Myh7* locus reveals open chromatin peaks in both the left atrial and cardiomyocyte ATAC-seq datasets (Figure 3.S3A). We also generated a fibroblast

ATAC-seq dataset using a differential adhesion method to isolate the fibroblast cell population [198]. Examination of the *Tcf21* locus, a fibroblast-specific gene, reveals open chromatin peaks in the left atria ATAC-seq and fibroblasts ATAC-seq datasets. The open chromatin peaks in all three ATAC-seq datasets are enriched in the promoter, intronic, and intergenic regions of the genome (Figure 3.S3C-D). To examine the quality of the called peaks we compared our ATAC-seq datasets with an adult heart ENCODE DNase (ENCFF977BOJ), which resulted in around 70% overlap (cardiomyocyte 72%; fibroblast 67%; left atria 73%; Figure 3.S3E). Direct comparisons of the cardiomyocyte and fibroblast ATAC-seq datasets results in cell type specific peaks, and gene ontology analysis of their promoters identified terms such as "regulation of heart contraction (GO:0008016)" and "chemotaxis (GO:0006935)" (Figure 3.S3F-H). An overlap of the cell-specific ATAC-seq datasets and differentially expressed ncRNA transcripts identifies candidate regulatory elements that are controlling gene expression in homeostasis and cardiac disease (Figure 3.2J-L). Overlap of the left atria ATAC-seq and DE ncRNAs identified 1,508 TAC-dependent enhancers and 1,042 *Tbx5*-dependent enhancers, with 468 of these regulatory elements shared in atrial fibrillation and pressure overload (Figure 3.2J).

We subsequently wanted to identify potential transcription factors regulating the expression of these candidate regulatory elements by performing motif analysis. For the candidate enhancers at down-regulated ncRNAs in the TBX5 KO model we detect an enrichment of canonical cardiac transcription factors like T-BOX, MEF, and GATA (Figure 3.2F). Motif analysis of the enhancers at the down-regulated ncRNAs in the TAC model share many of the same motifs, including the canonical cardiac transcription factors (Figure 3.2G). When we independently examine the motif enrichment of the enhancers from the upregulated ncRNAs in TBX5 KO and TAC mouse models, we see an enrichment for similar transcription factors including JUN/FOS,

STAT, and ETS. These candidate transcription factors are early response genes or have previously been implicated in disease processes, therefore making them interesting candidate regulators of cardiac disease.

Differential deep sequencing of ncRNAs identifies a TBX5-dependent enhancer downstream of Klf15.

Our laboratory has previously focused on the *Tbx5*-dependent down-regulated ncRNA transcripts that identified regulatory elements for ion channels and calcium handling genes, which was instrumental in understanding the TBX5 requirement for calcium homeostasis. Our analysis has identified a gene regulatory network that is shared in atrial fibrillation and pressure overload, and we have since confirmed many of these ncRNA transcripts are shared in the two mouse models, including RACER at the *Ryr2* locus as shown in Figure 3.2C. Closer examination of the differentially expressed genes near the putative regulatory elements revealed a candidate enhancer downstream of Kruppel-like factor 15 (*Klf15),* one of the shared transcription factors that has decreased expression in the mouse models (Figure 3.1K). *Klf15* is a particularly interesting candidate because it is known to have important roles in cardiac pathophysiology, where previous studies have revealed that it is a repressor of cardiac hypertrophy and fibrosis [206]–[209]. As shown in the genome browser view, there is decreased expression of the ncRNA and *Klf15* gene in both atrial fibrillation and pressure overload (Figure 3.3A). The ncRNA transcription is located 31 kb downstream of *Klf15*, and bordering the ncRNA are three adjacent open chromatin peaks in the left atria, cardiomyocyte, and fibroblast ATAC-seq datasets. These open chromatin regions align with peaks in several publicly available adult mouse heart datasets, including DNase-seq (ENCODE-ENCFF977BOJ), H3K27ac ChIP-seq (ENCODE-ENCFF445IUE), H3K4me1 ChIP-seq (ENCODE-ENCFF163PGG), POLR2A ChIP-seq (ENCODE-ENCFF454MFZ), and p300

ChIP-seq (ENCODE-ENCFF383PSV); all of which are known to be associated with enhancer activity (Figure 3.3A). DNase-seq is a genomic datasets that is sensitive to DNase I cleavage, which can be utilized to identify open chromatin regions in promoter and enhancers. H3K27ac and H3K4me1 are histone marks that are utilized to identify active transcription and enhancers, and polymerase II and p300 are enzymes involved in transcriptional processes. These ENCODE datasets are canonical markers for enhancer activity, therefore implicating our region of interest as an active enhancer.

In order to determine the gene that is regulated by our candidate enhancer, we explored the promoter-enhancer interactions utilizing a micro-C dataset generated by our collaborators in the Ruthenburg laboratory at the University of Chicago. Micro-C utilized chromatin fragmentation with micrococcal nucleases to generate a three-dimensional view of genomic structures to generate high resolution contact maps. While *Klf15* is the most proximal gene to the candidate enhancer, the micro-C dataset confirms there is an interaction between our candidate enhancer and *Klf15*. This interaction provides evidence that this region is an enhancer for *Klf15* gene regulation.

Transcription at the *Klf15* gene and enhancer are decreased after the removal of *Tbx5*. To elucidate whether this enhancer is directly regulated by TBX5, we generated a TBX5 ChIP-seq dataset from the adult heart tissue. Our candidate enhancer has TBX5 ChIP-seq binding sites at the promoter and downstream enhancer of *Klf15*, which suggests that this transcription factor plays an important role in *Klf15* gene regulation (Figure 3.3A). A publicly available TBX5 ChIP-seq dataset in the adult mouse ventricle also demonstrates TBX5 binding at these regions, which validates the functional role of TBX5 in regulating *Klf15* expression [58]. This is a fascinating finding because it implies TBX5 is not only important for regulating ion handling genes, but also it is an important regulator of cardiac homeostasis by repressing hypertrophy.

The *Klf15* enhancer consists of three adjacent ATAC-seq peaks that are ~3,500 basepairs total. To experimentally validate whether these open chromatin regions are active enhancers in cardiomyocytes, they were evaluated in an *in vitro* luciferase reporter assay. The three ATAC-seq peaks (Peak 1, Peak 2, and Peak 3) were cloned into luciferase expression plasmids, as shown in Figure 3.3B. All three candidate enhancers demonstrated robust reporter activity in HL-1 cardiomyocytes, indicative of cardiac enhancer function (Figure 3.3C; Peak 1: Fold Change=6.4, p-value=0.027; Peak 2: fold change=40.8, p-value=0.003; Peak 3: Fold Change=17.1, p-value=0.04). To experimentally validate whether TBX5 can regulate these enhancers, a luciferase reporter assay was performed in HEK293T cells with TBX5 overexpression. These enhancers were strongly TBX5-dependent, as the activity of these enhancers was significantly increased with the overexpression of the protein [Figure 3.3D; Peak 1: Fold Change=0.87, p-value=0.61; Peak 1 (with TBX5): Fold Change=5.79, p-value=0.03; Peak 2: Fold Change=0.76, p-value=0.27; Peak 2 (with TBX5): Fold Change=8.64, p-value=0.02; Peak 3: Fold Change=1.41, p-value=0.36; Peak 3 (with TBX5): Fold Change=7.0, p-value=0.02]. Therefore, we have identified an enhancer for *Klf15* that is directly regulated by TBX5.

We next interrogated whether this *Klf15* regulatory element has relevance to human disease. Our collaborators in the laboratory of Igor Efimov generated a Cap Analysis Gene Expression sequencing (CAGE-seq) dataset in human healthy and failing hearts to identify promoter and enhancer regions relevant to cardiac disease. The CAGE-seq peaks were lifted over from hg38 in humans to mm10 in mice, and a CAGE-seq peak was identified at our candidate *Klf15* enhancer of interest (Figure 3.3A). We examined the tracks for this locus in the human CAGE-seq dataset and determined that the peak is down-regulated in failing atria and the failing ventricle (Figure 3.3E). Adjacent to these CAGE-seq peaks are regions of open chromatin in

human ATAC-seq and DNase-seq datasets. The open chromatin regions are also enriched with H2K27ac and H3K4me1 ChIP-seq peaks, indicative of an active enhancer. To further characterize this enhancer in the human context, we examined the locus in the publicly available promoter capture Hi-C (PCHi-C) maps in iPSC-derived cardiomyocytes from the laboratory of Marcelo Nobrega (Figure 3.3F) [210]. This dataset shows an interaction between our lifted over candidate enhancer and the *Klf15* gene, very similar to what we see in the mouse. Therefore, our analysis in the human heart failure dataset demonstrates that the ncRNA-enhancer region in the mouse has relevance to human disease.

Conserved lncRNA is upregulated in both atrial fibrillation and heart failure mouse model.

Most models for transcriptional dysregulation presume that perturbation of a wild-type gene regulatory network as the cause of a disease risk, but examining the upregulated transcripts after a perturbation implicates the activation of a disease-specific gene regulatory network. Therefore, while we have examined several down-regulated ncRNA transcripts and their adjacent enhancers, it was important to also explore the up-regulated ncRNA transcripts. The up-regulated ncRNA transcripts are activated in a disease context, and therefore we hypothesized that the adjacent regulatory elements may reveal disease-responsive enhancers that are essential for coping with atrial dysfunction. Examining the differentially expressed genes proximal to up-regulated ncRNA transcripts revealed an interesting candidate upstream of SRY-Box Transcription Factor 9 (*Sox9*). This transcription factor is one of the shared candidate transcription factors that has increased expression in both *Tbx5* KO and TAC (Figure 3.1K). *Sox9* is particularly interesting because it is already a known regulator of fibrosis processes [211]–[213]. The ncRNA transcript is ~1,000 kb away from the *Sox9* promoter, and the transcript is up-regulated in the *Tbx5* KO and TAC samples. At the 5' end of the ncRNA there is a left atria ATAC-seq peak and a DNase peak

(ENCODE-ENCFF977BOJ), but there is an absence of active enhancer marks including H3K27ac (shown in Figure 3.3A, ENCODE-ENCFF445IUE), H3K4me1 (not shown, ENCODE-ENCFF163PGG), POLR2A (not shown, ENCODE-ENCFF454MFZ), and p300 (not shown, ENCODE-ENCFF383PSV). We hypothesize that this enhancer is activated in a disease-specific context, so it is not surprising there is an absence of active enhancer marks from the ENCODE wild-type datasets. The enhancer, characterized by the open chromatin peak at the 5' end of the ncRNA, does not have any looping interactions from the prospective of the cardiomyocyte micro-C dataset. Due to the proximity and differential expression of the gene and the ncRNA, we implicate this putative enhancer is a potential regulatory mechanism for *Sox9*.

The differentially expressed genes have implicated transcriptional changes in many cell types of the heart, including the fibroblasts. Due to the fact that *Sox9* is a fibroblast-specific gene, we believe this ncRNA is also expressed in the fibroblasts. This ncRNA is up-regulated in atrial fibrillation and pressure overload, therefore the transcript is activated in cardiac disease contexts. During cardiac remodeling processes, the fibroblasts transition from a quiescent phenotype to a pathological phenotype where the fibroblasts excrete an increased amount of collagens and extracellular matrix proteins. TGF-β treatment is a widely accepted method of activating fibroblasts in an *in vitro* context [134], [214]. Primary fibroblasts were isolated from the adult mouse heart and treated with TGF-β in order to evaluate the regulatory mechanisms of cardiac fibroblasts in a disease context (Figure 3.4B). Transcriptional profiling on the control primary fibroblasts (vehicle) and TGF-β treated primary fibroblasts revealed 2,665 differentially expressed genes (Figure 3.S4A; 1,489 genes, log2FoldChange<0.5, p.adj<0.05; 1,176 genes, log2FoldChange>0.5, p.adj<0.05). Gene ontology analysis of the differentially expressed genes revealed an enrichment for terms related to "extracellular matrix organization (GO:0030198)" and

"collagen formation (R-MMU-1474290)" (Figure 3.S4A). The enrichment of these terms is expected as TGF-β treatment is eliciting fibroblast activation. The overlap of the differentially expressed genes in TGF-β treated fibroblasts, TAC, and *Tbx5*-mutant experiments, identified shared transcriptional changes between the cardiac perturbation mouse models and *in vitro* fibroblast activation (Figure 3.S4A-B). A heatmap of the shared genes displayed the directionality of changed gene expression, and revealed an abundance of genes that were both shared and unique compared to the mouse models (Figure 3.S4C).

The treatment of fibroblasts with TGF-β also leads to changes in the chromatin landscape, which revealed 12,130 differentially open chromatin peaks in an ATAC-seq comparison of the vehicle vs. TGF-β treated fibroblasts (Figure 3.4D, Figure 3.S4D; 4,566 closing peaks, log2FoldChange<0.5, FDR<0.2; 7,564 opening peaks, log2FoldChange>0.5, FDR<0.2). To elucidate the potential regulators of the changing chromatin landscape, motif analysis was performed on the opening and closing ATAC-seq peaks. For the closing chromatin peaks there is an enrichment of motifs for CEBP, KLF, and STAT family of proteins. The opening chromatin peaks have an enrichment for motifs including FRA, SMAD, TCF, and ETS family of motifs. It is reassuring to see the presence of the SMAD motif as this family of proteins is induced in response to TGF-β. The identification of these transcription factors may identify transcriptional regulators of fibroblast activation, which may provide insight into fibrotic processes.

To elucidate the regulatory mechanisms driving fibroblast activation, we identified regulatory elements by identifying the differentially expressed ncRNA transcripts. The treatment of fibroblasts with TGF-β led to 2,594 significantly differentially expressed ncRNA transcripts. A closer inspection of our candidate ncRNA downstream of *Sox9* reveals it is expressed in the TGF-

β treated fibroblasts. We have now demonstrated that this ncRNA candidate is expressed in the fibroblast population and is activated in a disease state. Open chromatin accessibility and differential transcription of the ncRNA transcript implicated Peak 4 as an active enhancer. Closer examination of the chromatin accessibility in the locus demonstrates Peak 4 has an opening ATAC-seq peak, providing evidence for increased accessibility of this candidate enhancer following TGF-β treatment.

We wanted to examine whether ncRNA transcription could identify the most active enhancer in the locus. Five enhancers defined by the fibroblast ATAC-seq peaks and DNase peaks were cloned into luciferase reporter plasmids to evaluate enhancer activity in 3T3 fibroblasts. The candidate enhancer at the 5' end of the DE ncRNA (Peak 4) showed strong enhancer activity in the luciferase reporter assay in fibroblasts (Figure 3.4F; FC= 14.9, p-value= 1.39E-6). This is in contrast to the remaining candidates which had little to no enhancer activity in fibroblasts (Figure 3.4F). Our candidate ncRNA is activated in a disease context, we therefore wanted to examine whether the enhancer has more activation in a disease state as well. Therefore, we treated the fibroblasts with TGF-β in a luciferase assay, which showed an increased amount of activation (Figure 3.4F; FC= 29.9, p-value= 8.9E-6). The results of this analysis reveal we have identified a fibroblast ncRNA-borne enhancer that is activated during cardiac disease, including atrial fibrillation and pressure overload.

While our previous down-regulated candidates have been regulated by TBX5 expression, the up-regulated candidates are activated in the absence of this transcription factor. In an effort to identify the transcription factors that are driving the activity of the enhancer, we examined the motifs that were shown in the opening ATAC-seq peaks shown in Figure 3.4D. We focused our interests on SMAD and TCF as these transcription factors are expressed in fibroblasts. The TGF-

β signal transduction pathway results in accumulation and activation of SMAD proteins that directly regulate downstream gene expression, which makes this an interesting candidate for the activation of our ncRNA with TGF-β treatment [214], [215]. TCF21 is expressed in cardiac fibroblast and is crucial for the development of fibroblasts, as they fail to form in the absence of TCF21 [216]. Therefore, we independently mutated the binding sites of SMAD and TCF family of transcription factors in the enhancer at Peak 4. The mutation of TCF binding sites in Peak 4 resulted in no luciferase activation, implicating this as an important transcription factor in the regulation of this enhancer. Interestingly, the mutation of the SMAD binding sites resulted in increased activation of the candidate enhancer. This suggests that enhancer is regulated by SMAD proteins, but further analysis is required to determine its exact function in the regulation of the enhancer.

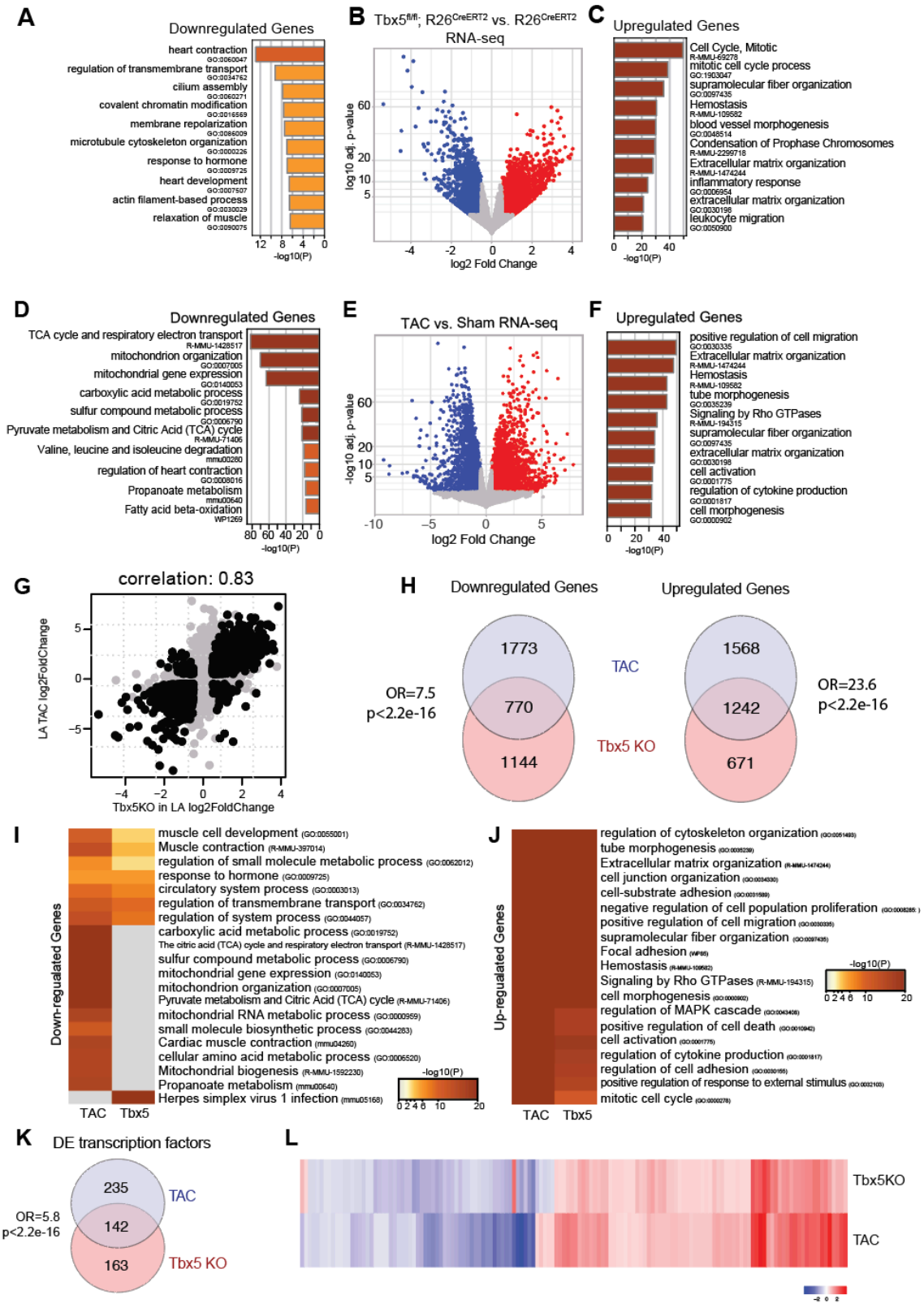


Figure 3.1. Conserved transcriptional networks in atrial fibrillation and heart failure mouse model.

Figure 3.1, continued

A-C. RNA-seq analysis in left atria of *R26*^{CreERT2} (control) vs. *Tbx5*$^{fl/fl}$,*R26*^{CreERT2} (*Tbx5* KO). (A) GO term analysis on down-regulated genes from control and *Tbx5* KO comparison (sorted by -log10(p-value)). (B) Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of control vs. *Tbx5* KO coding transcripts in RNA-seq. Significantly downregulated genes (log2FoldChange<0.5 and adj. p-value<0.05; 1914 genes) are shown in blue. Significantly upregulated genes (log2FoldChange>0.5 and adj. p-value<0.05; 1913 genes) are shown in red. Non-significant differentially expressed transcripts are grey. (C) GO term analysis on significant up-regulated genes from *Tbx5* KO and control comparison (sorted by - log10(p-value)). D-F. RNA-seq analysis in left atria of sham vs. transverse aortic constriction (TAC) mice. (D) GO term analysis on down-regulated genes from sham and TAC comparison (sorted by - log10(p-value)). (E) Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of sham vs. TAC coding transcripts in RNA-seq. Significantly downregulated genes (log2FoldChange<0.5 and adj. p-value<0.05; 1926 genes) are shown in blue. Significantly upregulated genes (log2FoldChange>0.5 and adj. p-value<0.05; 1931 genes) are shown in red. Non-significant differentially expressed transcripts are grey. (F) GO term analysis on significant up-regulated genes from sham and TAC comparison (sorted by -log10(p-value)). G. Scatterplot for genes identified in the control vs. *Tbx5* KO (log2FoldChange, x-axis) and sham vs. TAC (log2FoldChange, y-axis) RNA-seq analysis. Light grey dots represent all identified genes, while black dots represent significantly differentially expressed genes (|log2FoldChange|>0.5. p. adj <0.05) in both comparisons. Pearson's correlation for comparison is 0.82. H. Venn diagram of TAC down-regulated genes and *Tbx5* KO down-regulated genes (left); venn diagram of TAC up-regulated genes and *Tbx5* KO up-regulated genes (right). I-J. Comparative analysis using Metascape for down-regulated genes (I) and up-regulated genes (J) identified in sham vs. TAC (left column) and control vs. *Tbx5* KO (right column). Metascape utilizes Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Heatmap of enriched terms are colored by -log10(p-value). K. Venn diagram of differentially expressed transcription factors (|log2FoldChange|>0.5; p. adj <0.05) in sham vs. TAC (blue) and control vs. *Tbx5* KO (red). L. Heatmap of log2FoldChange for 132 shared differentially expressed genes between control vs. *Tbx5* KO (top row) and sham vs. TAC (bottom row).

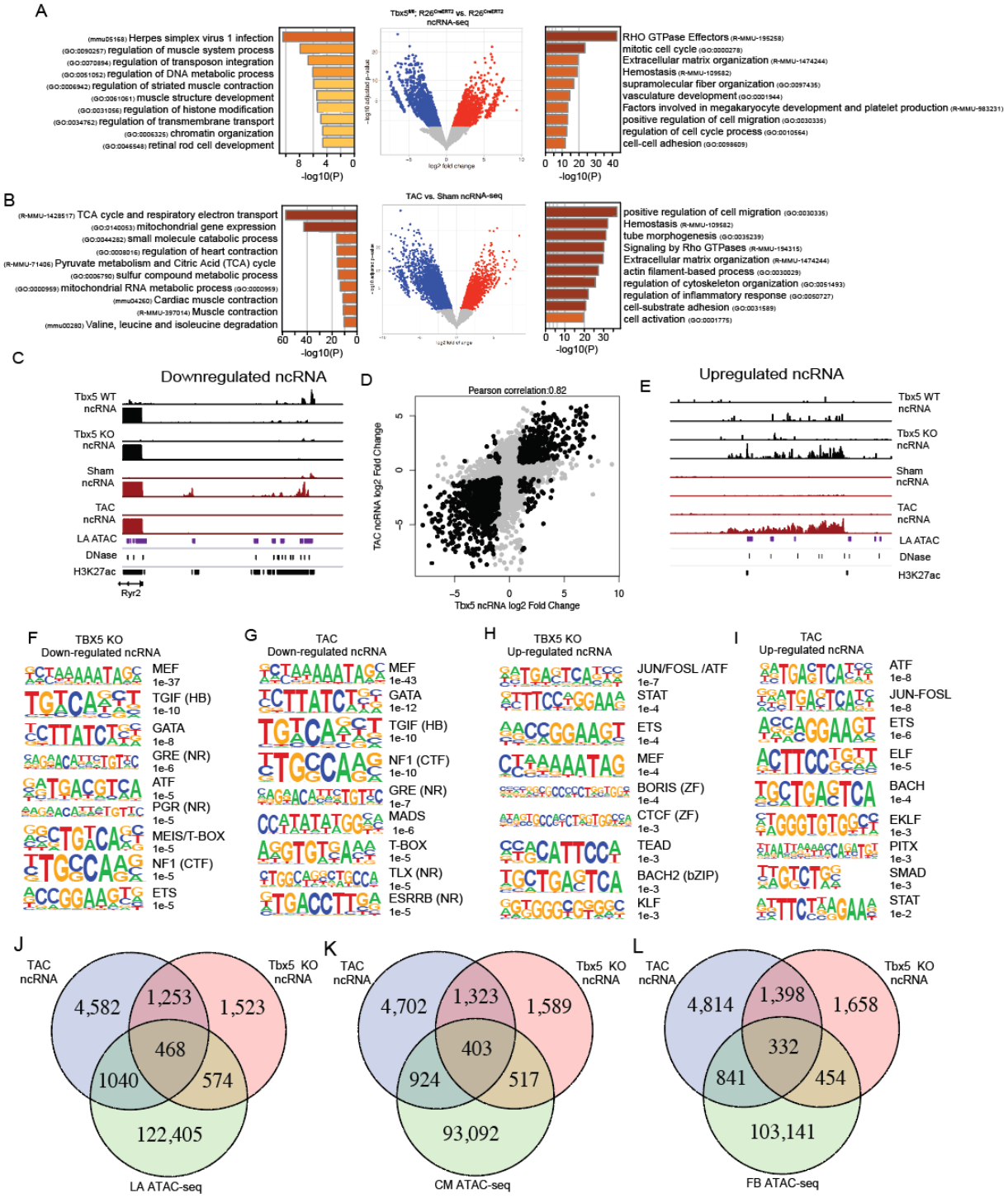


Figure 3.2. Positive correlation in shared ncRNAs implicates shared regulatory networks in atrial fibrillation and heart failure. A. ncRNA-seq analysis in left atria of *R26*^{CreERT2} (control) vs. *Tbx5*$^{fl/fl}$,*R26*^{CreERT2} (*Tbx5* KO). Left: GO term analysis on differentially expressed genes in the same TAD as down-regulated ncRNAs from control and *Tbx5* KO comparison (sorted by -log10(p-value)). Center: Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of control vs.

Figure 3.2, continued

Tbx5 KO transcripts in ncRNA-seq. Significantly downregulated ncRNA transcripts (log2FoldChange<0.5 and adj. p-value<0.05) are shown in blue. Significantly upregulated ncRNA transcripts (log2FoldChange>0.5 and adj. p-value<0.05) are shown in red. Non-significant transcripts are grey. Right: GO term analysis on differentially expressed genes in the same TAD as up-regulated ncRNA transcripts from *Tbx5* KO and control comparisons (sorted by - log10(p-value)). B. ncRNA-seq analysis in left atria of sham vs. transverse aortic constriction (TAC) mice. Left: GO term analysis on differentially expressed genes in the same TAD as down-regulated ncRNAs from sham and TAC comparison (sorted by - log10(p-value)). Center: Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of sham vs. TAC ncRNA transcripts in ncRNA-seq. Significantly downregulated ncRNA transcripts (log2FoldChange<0.5 and adj. p-value<0.05; 1926 genes) are shown in blue. Significantly up-regulated ncRNA transcripts (log2FoldChange>0.5 and adj. p-value<0.05; 1931 genes) are shown in red. Non-significant transcripts are depicted by grey dots. Right: GO term analysis on differentially expressed genes in the same TAD as up-regulated ncRNA transcripts from sham and TAC comparison (sorted by -log10(p-value)). C. Genomic view of a shared down-regulated ncRNA transcript. The tracks are as follows (mm10): *Tbx5* WT forward ncRNA-seq, *Tbx5* WT reverse ncRNA-seq, *Tbx5* KO forward ncRNA-seq, *Tbx5* KO reverse ncRNA-seq, sham forward ncRNA-seq, sham reverse ncRNA-seq, TAC forward ncRNA-seq, TAC reverse ncRNA-seq, left atrial ATAC-seq, ENCODE adult mouse heart DNase (ENCFF977BOJ), and ENCODE adult mouse H3K27ac ChIP-seq (ENCFF445IUE). Genomic location: chr13:12,104,827-12,143,786. D. Scatterplot for ncRNA transcripts identified in the control vs. *Tbx5* KO (log2FoldChange, x-axis) and sham vs. TAC (log2FoldChange, y-axis) ncRNA-seq. Light grey dots represent all identified ncRNA transcripts, while black dots represent significant differentially expressed ncRNA transcripts in both comparisons. Pearson's correlation for the comparison is 0.82. E. Genomic view of a shared up-regulated ncRNA transcript in the *Tbx5* KO and TAC mouse models. The tracks are as follows (mm10): *Tbx5* WT forward ncRNA-seq, *Tbx5* WT reverse ncRNA-seq, *Tbx5* KO forward ncRNA-seq, *Tbx5* KO reverse ncRNA-seq, sham forward ncRNA-seq, sham reverse ncRNA-seq, TAC forward ncRNA-seq, TAC reverse ncRNA-seq, left atrial ATAC-seq, ENCODE adult mouse heart DNase (ENCFF977BOJ), and ENCODE adult mouse H3K27ac ChIP-seq (ENCFF445IUE). Genomic location: chr2: 163,789,659-163,814,620. F-I. Identification of motifs for ncRNAs with a left atrial ATAC peak at the 5' end of the transcript. Motif analysis for down-regulated ncRNAs in the *Tbx5* KO (F) and TAC (G) mouse models. Motif analysis for up-regulated ncRNAs in the *Tbx5* KO (H) and TAC (I) mouse models. J. Venn diagram of differential ncRNA transcripts in TAC (blue) and differential ncRNA transcripts in *Tbx5* KO (red), overlapped with left atria (LA) ATAC peaks. K. Venn diagram of differential ncRNA transcripts in TAC (blue) and differential ncRNA transcripts in *Tbx5* KO (red), overlapped with cardiomyocyte (CM) ATAC peaks. L. Venn diagram of differential ncRNA transcripts in TAC (blue) and differential ncRNA transcripts in *Tbx5* KO (red), overlapped with fibroblast (FB) ATAC peaks.

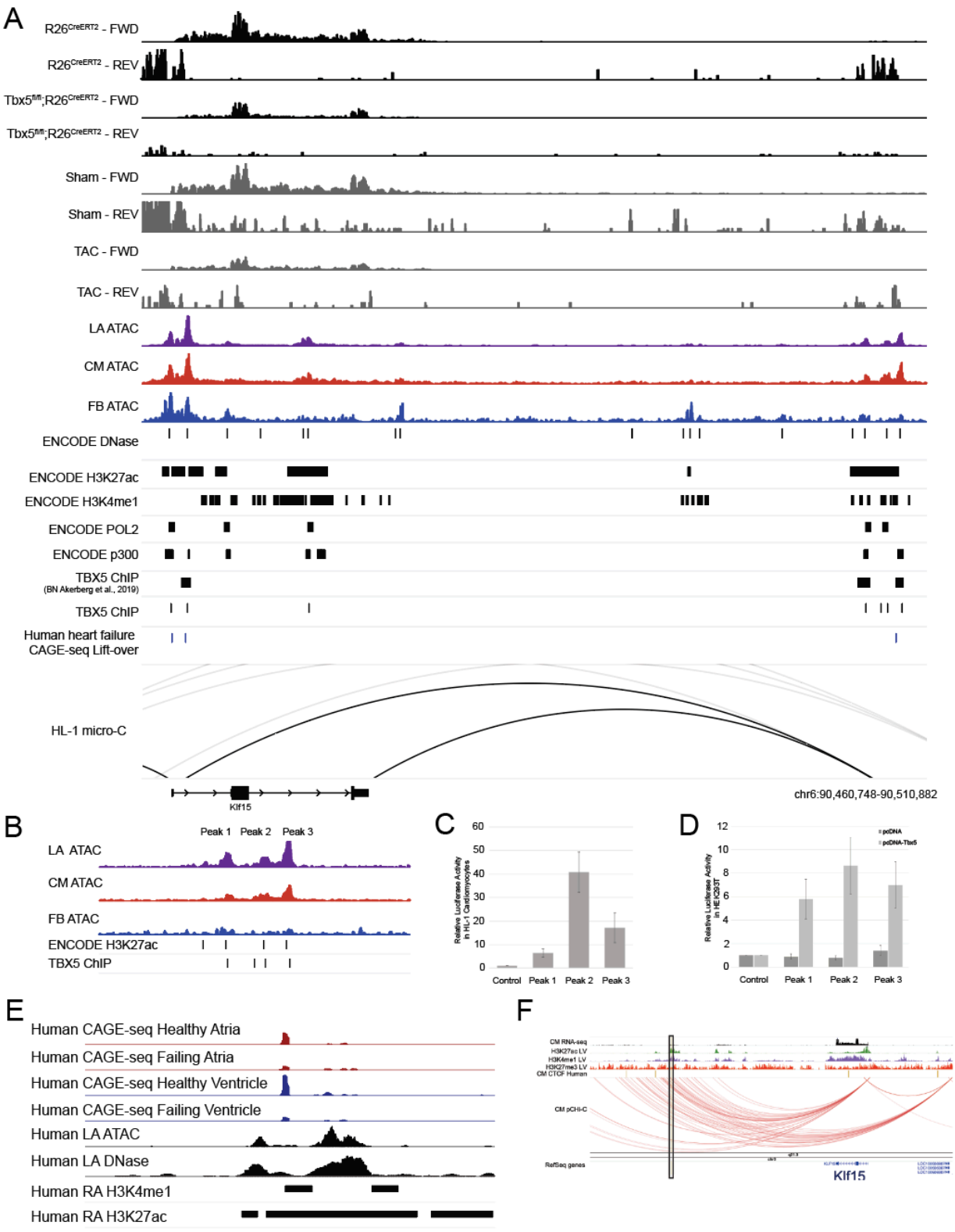


Figure 3.3. Differential deep sequencing of ncRNAs identifies TBX5-dependent enhancer downstream of *Klf15*. A. Genomic view of the *Klf15* locus with nearby downstream enhancer. Genomic tracks are as follow: control (*R26*^{CreERT2}) ncRNA-seq forward, control (*R26*^{CreERT2})

Figure 3.3, continued

ncRNA-seq reverse, *Tbx5* KO (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) ncRNA-seq forward, *Tbx5* KO (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) ncRNA-seq reverse, sham ncRNA-seq forward, sham ncRNA-seq reverse, TAC ncRNA-seq forward, TAC ncRNA-seq reverse, left atria (LA) ATAC-seq, cardiomyocyte (CM) ATAC-seq, fibroblast (FB) ATAC-seq, ENCODE adult mouse heart DNase (ENCFF977BOJ), ENCODE adult mouse heart H3K27ac (ENCFF445IUE), ENCODE adult mouse heart H3K4me1 (ENCFF163PGG), ENCODE adult mouse heart POLR2A ChIP-seq (ENCFF454MFZ), ENCODE adult mouse heart p300 (ENCFF383PSV), TBX5 ChIP-seq (Akerberg et al., 2019) [58], TBX5 ChIP-seq (in-house), Human heart failure CAGE-seq lifted over to mouse mm10, micro-C in HL-1 cardiomyocytes. B. Schematic depicting the genomic tracks for the ATAC peaks and candidate enhancers for *Klf15*. Genomic tracks are as follow: left atria (LA) ATAC-seq, cardiomyocyte (CM) ATAC-seq, fibroblasts (FB) ATAC-seq, ENCODE adult mouse heart H3K27ac (ENCFF445IUE), and TBX5 ChIP-seq. C-D. Luciferase reporter assays for the candidate *Klf15* regulatory elements. Peak 1, Peak 2, Peak 3 correspond to the ATAC peaks in Figure 3.3B. Data are shown as mean $\pm$ SEM normalized to a control vector. P-values were determined with a two-tailed t test. Luciferase assay was performed in HL-1 cardiomyocytes (C), and HEK293T cells transfected with TBX5 expression vector (D). E. Genomic view of candidate enhancer downstream of *Klf15* in humans with the tracks as follows: human CAGE-seq healthy atria, human CAGE-seq failing atria, human CAGE-seq healthy ventricle, human CAGE-seq failing ventricle, human left atria (LA) ATAC (ENCODE ENCFF631IWY), human left atria (LA) DNase (ENCODE ENCFF217MSL), human right atria (RA) H3K4me1 (ENCODE ENCFF409XGI), human right atria (RA) H3K27ac (ENCODE ENCFF153OUB). F. Genomic view iPSC-induced cardiomyocyte Hi-C interaction at the *Klf15* locus. Genomic tracks as follows: cardiomyocyte RNA-seq, left ventricle H3K27ac (GSM908951), left ventricle H3K4me1 (GSM910575), left ventricle H3K27me3 (GSM908952), CTCF human cardiomyocytes, iPSC-induced cardiomyocyte Hi-C [210].

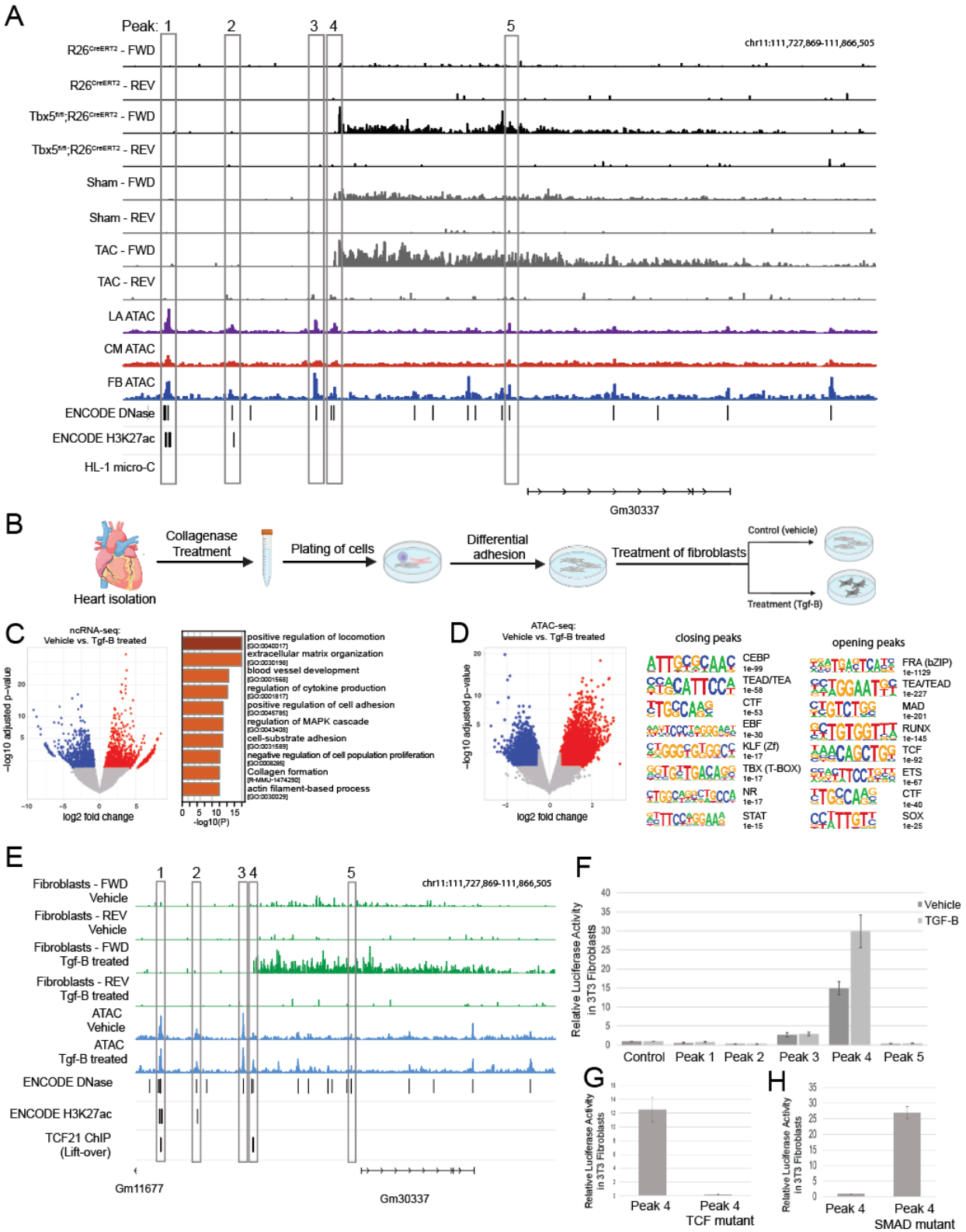


Figure 3.4. Conserved ncRNA is upregulated in both atrial fibrillation and heart failure mouse model. A. Genomic view of shared ncRNA with genomic tracks as follow: control (*R26*$^{\text{CreERT2}}$) ncRNA-seq forward, control (*R26*$^{\text{CreERT2}}$) ncRNA-seq reverse, *Tbx5* KO (*Tbx5*$^{\text{fl/fl}}$;*R26*$^{\text{CreERT2}}$)

Figure 3.4, continued

ncRNA-seq forward, *Tbx5* KO (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) ncRNA-seq reverse, sham ncRNA-seq forward, sham ncRNA-seq reverse, TAC ncRNA-seq forward, TAC ncRNA-seq reverse, left atria (LA) ATAC-seq, cardiomyocyte (CM) ATAC-seq, fibroblast (FB) ATAC-seq, ENCODE adult mouse heart DNase (ENCFF977BOJ), ENCODE adult mouse heart H3K27ac (ENCFF445IUE), and micro-C in HL-1 cardiomyocytes. B. Schematic of isolation and treatment of primary cardiac fibroblasts. C. ncRNA-seq analysis of cultured primary fibroblasts in control (vehicle) and activated (TGF-β treated) conditions. Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of control vs. TGF-β treated fibroblasts in ncRNA-seq. Significantly downregulated transcripts (log2FoldChange<0.5 and adj. p-value<0.05) are shown in blue. Significantly upregulated transcripts (log2FoldChange>0.5 and adj. p-value<0.05) are shown in red. Non-significant differentially expressed transcripts are grey. GO term analysis on significant differentially expressed ncRNAs within 2 megabases of differentially expressed genes, sorted by -log10(p-value). D. ATAC-seq analysis of cultured primary fibroblasts (control) vs. cultured primary fibroblasts treated with TGF-β (treatment). Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of differential ATAC-seq peaks comparing vehicle vs. TGF-β treated fibroblasts. Significant closing chromatin peaks (log2FoldChange<0.5 and adj. p-value<0.05) are shown in blue. Significant opening chromatin peaks (log2FoldChange>0.5 and adj. p-value<0.05) are shown in red. Non-significant transcripts are grey. Motif analysis of opening and closing chromatin peaks (right). E. Genomic view of shared lncRNA present in TGF-β treated fibroblasts, with tracks as follows: isolated fibroblast [vehicle] ncRNA-seq forward, isolated fibroblast [vehicle] ncRNA-seq reverse, isolated fibroblast treated with TGF-β ncRNA-seq forward, isolated fibroblast treated with TGF-β-ncRNA-seq reverse, isolated fibroblast [vehicle] ATAC-seq, isolated fibroblast treated with TGF-β ATAC-seq, ENCODE adult mouse heart DNase (ENCFF977BOJ), ENCODE adult mouse heart H3K27ac (ENCFF445IUE), lift over TCF21 ChIP-seq (Sazonova et al., 2015) [217]. F-H. Luciferase assays for candidate enhancers in the locus of the shared ncRNA. Data are shown as mean $\pm$ SEM normalized to a control vector. P-values were determined with a two-tailed t test. Luciferase assay was performed in 3T3 fibroblasts.

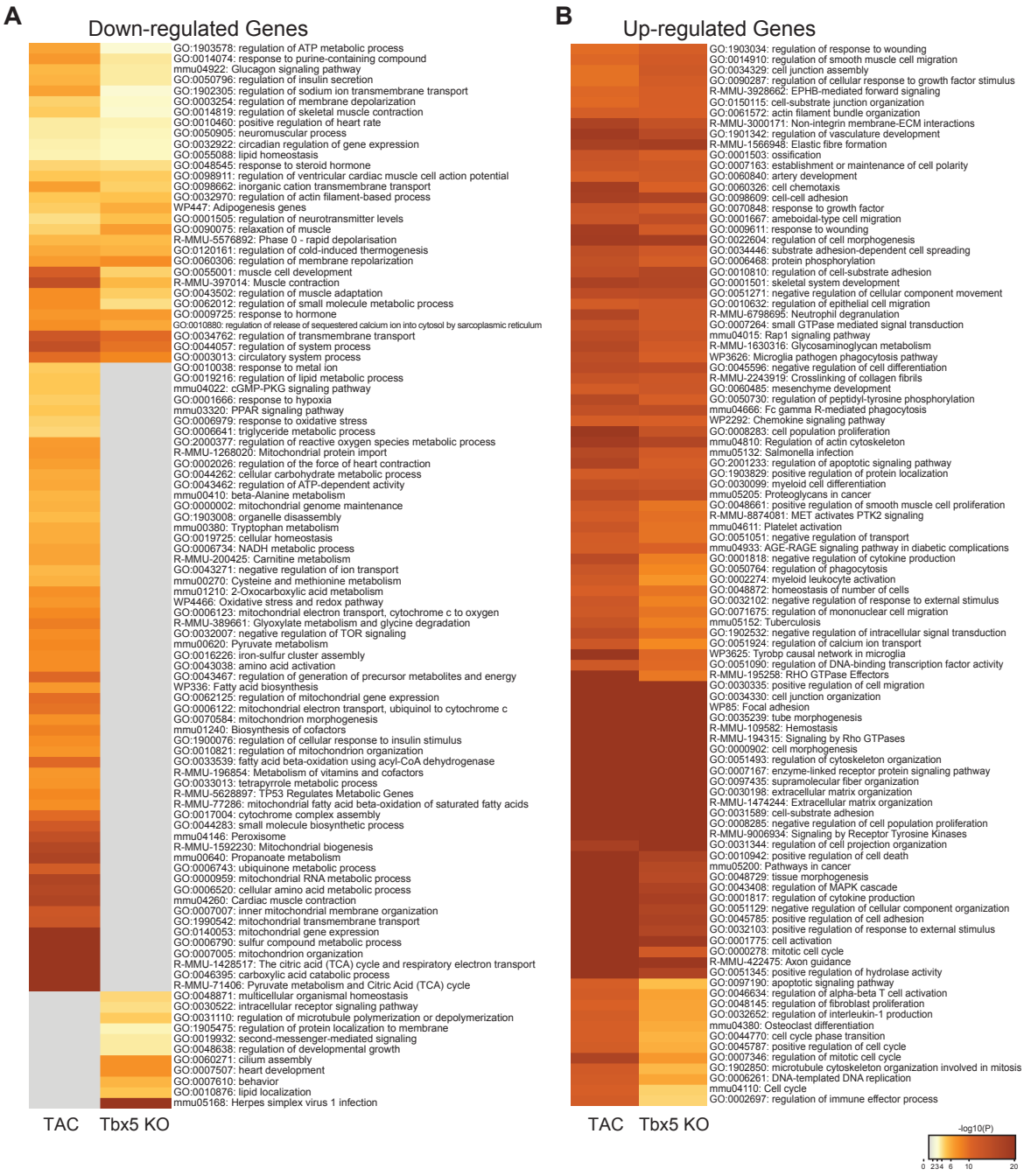


Figure 3.S1. Expanded comparative analysis for differential gene expression in atrial fibrillation and heart failure mouse model. Comparative analysis using Metascape for down-regulated genes (A) and up-regulated genes (B) identified in sham vs. TAC (left column) and control vs. *Tbx5* KO (right column). Heatmap of enriched terms are colored by -log10(p-value).

Figure 3.S2. Comparative pathway analysis in atrial fibrillation and heart failure mouse model. A-B. Ingenuity Pathway Analysis (IPA) analysis identified predicted repressed pathways (A) and predicted activated pathways (B) identified in sham vs. TAC (left column) and control vs. *Tbx5* KO (right column).

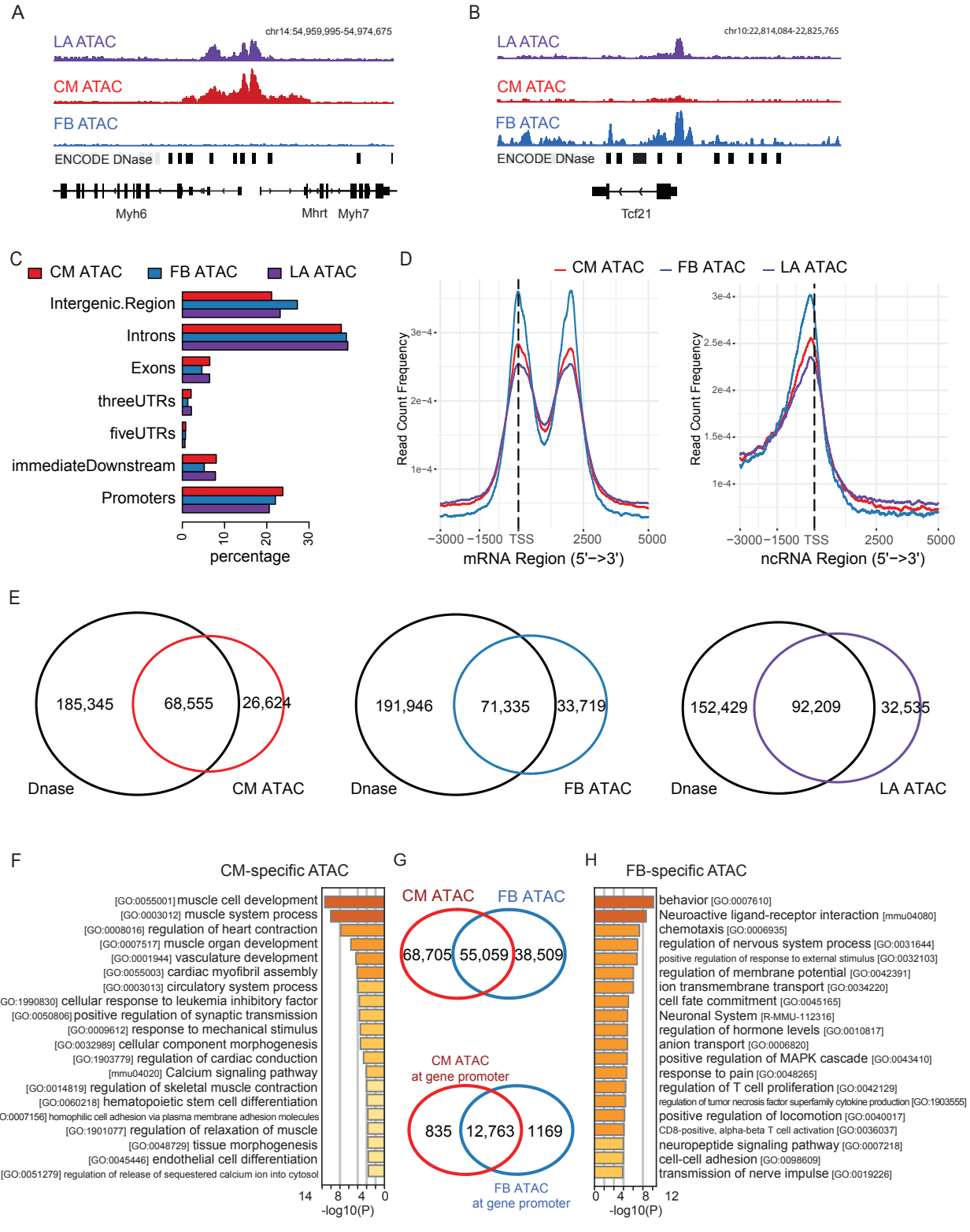


Figure 3.S3. Generation of cell-type specific ATAC-seq datasets in the mouse heart.

Figure 3.S3, continued

A-B. Genomic view of *Myh6-Myh7* locus (A) and *Tcf21* locus (B) with the following tracks: left atria (LA) ATAC-seq, cardiomyocyte (CM) ATAC-seq, fibroblast (FB) ATAC-seq, ENCODE adult heart DNase (ENCFF977BOJ). *Myh6* is a cardiomyocyte-specific gene, and Tcf21 is a fibroblast-specific gene, with a peak present at the gene promoter for each respective dataset. C. Bar graph depicting genomic features of peaks for the three ATAC-seq datasets: cardiomyocyte (red), fibroblast (blue), and left atria (purple). D. Visualization of read count frequency of peaks at the TSS of genes (left) and ncRNAs (right) for the three ATAC-seq datasets (CM, FB, LA). E. Venn diagram of ENCODE adult heart DNase peaks (union of ENCFF836BYB and ENCFF200YPA) with the CM ATAC-seq (left), FB ATAC-seq (middle), LA ATAC-seq (right). F-H. Comparison of cardiomyocyte ATAC-seq and fibroblast ATAC-seq datasets. (F) Gene ontology analysis for genes with CM-specific peaks (68,705 peaks, 835 gene promoters with a CM ATAC peak). (G) Venn diagram comparing cardiomyocyte and fibroblast ATAC peaks (top). Venn diagram depicting the number of gene promoters with cell-type specific ATAC peaks (bottom). (H) Gene ontology analysis for genes with FB-specific peaks (38,509 peaks, 1169 gene promoters with a FB ATAC peak).

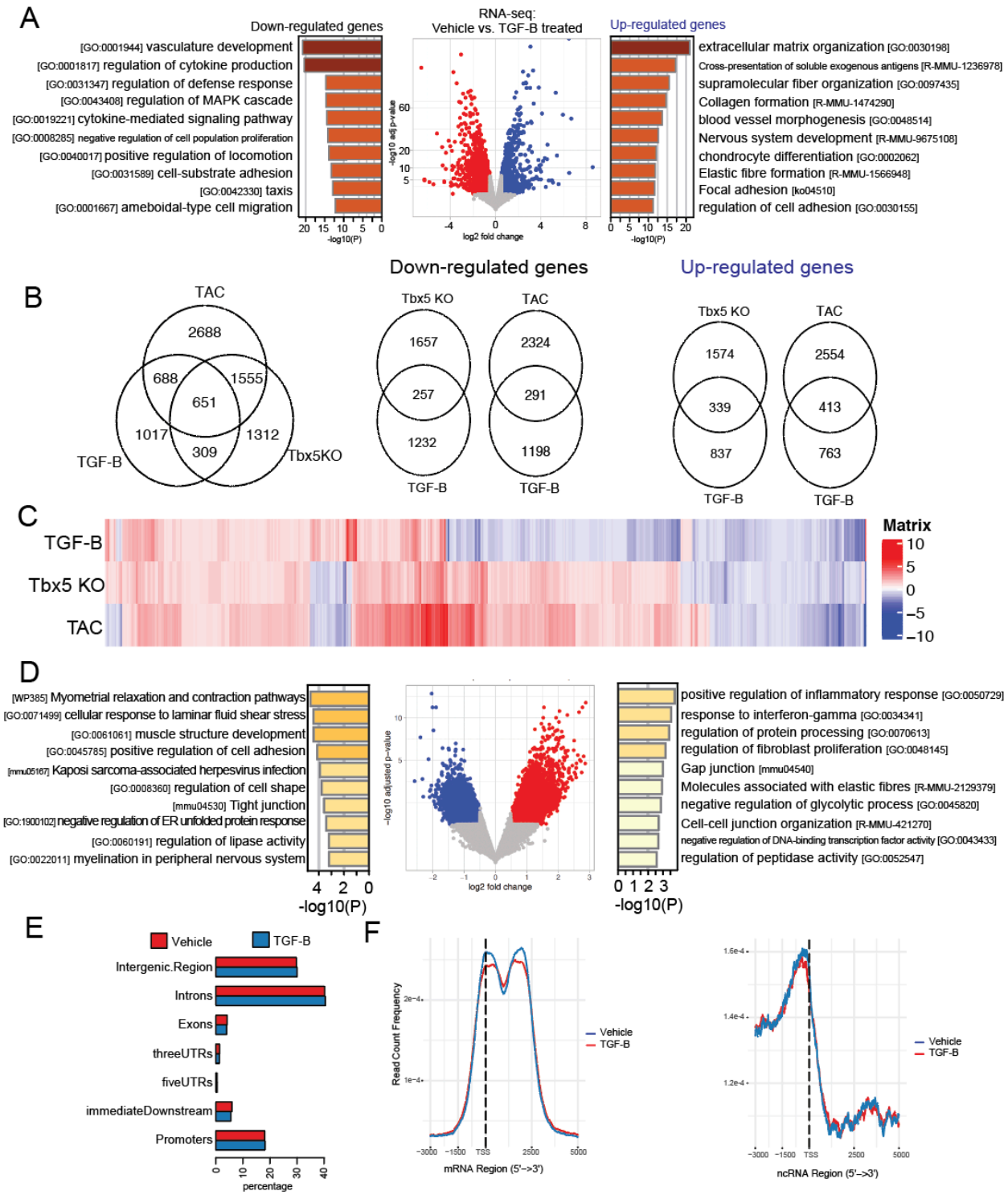


Figure 3.S4. Analysis of changes in gene expression and open chromatin landscape in activated fibroblasts. A-C. RNA-seq analysis of cultured primary fibroblasts in control (vehicle) and activated (TGF-β treatment) conditions. GO term analysis on significant down-regulated genes, sorted by -log10(p-value) (left). Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of control vs. TGF-β treated fibroblasts in RNA-seq. Significantly downregulated genes (log2FoldChange

Figure 3.S4, continued

<0.5 and adj. p-value<0.05; 1,489 genes) are shown in red. Significantly upregulated genes (log2FoldChange>0.5 and adj. p-value<0.05; 1,176 genes) are shown in blue. Non-significant differentially expressed transcripts are grey (middle). GO term analysis on significant up-regulated genes, sorted by -log10(p-value) (right). B. Venn diagram overlapping differentially expressed genes between TBX5 KO, TAC, and TGF-β treated fibroblasts (left). Venn diagram overlapping down-regulated genes between *Tbx5* KO and TGF-β treated fibroblasts; Venn diagram overlapping down-regulated genes between TAC and TGF-β treated fibroblasts (middle). Venn diagram overlapping up-regulated genes between *Tbx5* KO and TGF-β treated fibroblasts; Venn diagram overlapping up-regulated genes between TAC and TGF-β treated fibroblasts (right). C. Heatmap of the shared differentially expressed genes in the *Tbx5* KO, TAC, and TGF-β treated fibroblasts. D. ATAC-seq analysis in left atria of cultured primary fibroblasts (control) vs. cultured primary fibroblasts treated with TGF-β (treatment). GO term analysis on differentially expressed genes with closing chromatin peaks (left). Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of differential ATAC-seq peaks comparing vehicle vs. TGF-β treated fibroblasts (middle). Significantly closing chromatin peaks (log2FoldChange<0.5 and adj. p-value<0.05) are shown in blue. Significantly opening chromatin peaks (log2FoldChange>0.5 and adj. p-value<0.05) are shown in red. Non-significant transcripts are grey. GO term analysis on differentially expressed genes with opening chromatin peaks (right). E. Bar graph depicting the localization of chromatin peaks identified in ATAC-seq. F. Read count frequency of ATAC-seq peaks relative to mRNA promoters (left) and ncRNA promoters (right).

Chapter 3.4 Discussion

The strong epidemiologic link between heart failure atrial fibrillation prompted us to compare the gene regulatory networks of our atrial fibrillation and pressure overload induced heart failure mouse models. We examined the transcriptional changes occurring in the left atria of the *Tbx5*-deletion mouse model and compared it to the transcriptional changes in the left atria of the Transverse Aortic Constriction (TAC) mouse model. We revealed remarkable correlation between the differentially expressed genes between these distinct disease models, which suggests a shared regulatory network in atrial fibrillation and heart failure. This is a particularly interesting finding because the *Tbx5*-deficient samples were collected prior to the onset of atrial fibrillation, and well in advance of the cardiac remodeling that will occur several weeks later. This is in contrast to the TAC samples that were collected 2-weeks post-surgery, a timepoint at which the cardiac remodeling in the heart is well underway. This suggests that in atrial fibrillation the transcriptional changes for cardiac remodeling occurred much earlier than expected, and well in advance of visualization of the cardiac remodeling processes, including fibrosis. This is an exciting finding that has been seen before in atrial fibrillation progression studies in sheep [218]. Interestingly, this study found that transcriptional changes occurred early in the transition from paroxysmal to persistent atrial fibrillation, but even animals that were in persistent atrial fibrillation for another year did not have any distinct transcriptional changes. It is fascinating that the early transcriptional changes in atrial fibrillation have a shared transcriptional network with a heart failure model further down the remodeling process. A comparative gene ontology analysis of the up-regulated genes in the *Tbx5*-deficient and TAC RNA-seq revealed an enrichment for shared terms related to disease processes, while the comparison of the down-regulated terms revealed unique terms relevant to metabolism in the TAC model. The unique metabolic terms in the TAC mouse model are indicative

of canonical metabolic changes that occur in heart failure with the decreased production of high-energy phosphates and increased metabolic inefficacy in the heart [202], [203]. While our analysis shows the majority of the metabolic terms are unique to the TAC model, there is a shared "regulation of ATP metabolic process (GO:1903578)" term in the expanded gene ontology analysis. A closer examination of genes involved in metabolic processes may provide insight into any compensatory metabolic processes involved in atrial fibrillation.

Comparisons of differentially expressed transcription factors in the two cardiac diseases identified 142 candidates, 135 of which are changed in the same direction. We hypothesize these transcription factors may be important in driving the pathophysiology of cardiac diseases. One of the candidates identified in our analysis is Meox1 (Mesenchyme Homeobox 1), which was up-regulated in both the *Tbx5*-definicent and TAC mouse models (*Tbx5*-deletion log2FoldChange 1.61, p-adj. 2.63E-12; TAC log2FoldChage 2.16; p-adj. 8.29E-4). Meox1 and Meox2 (Mesenchyme Homeobox 2) are a pair of transcription factors from the homeobox protein family. Meox1 and Meox2 are crucial for somatogenesis during development, and in their absence there are defects in specification and patterning of the somites [219]. Recently, Meox1 has been shown to have a functional role in the adult heart as a crucial regulator of fibroblast activation in disease [134]. It is also confirmed to be a stress response gene in human disease as it is upregulated in human cardiac tissue from patients with cardiomyopathy and lung tissue from patients with pulmonary fibrosis [134]. A second example of a shared transcription factor is Runx1 (runt-related transcription factor 1), which was upregulated in both cardiac disease (*Tbx5*-deletion log2FoldChange 0.89, p-adj. 3.1E-5; TAC log2FoldChage 3.0; p-adj. 4.4E-4). Runx1 is well characterized as an important protein in the hematopoiesis and its altered expression can lead to myeloid leukemia, but it recent years its up-regulation in cardiac pathology is of particular interest.

The increased expression of *Runx1* has been seen in a variety of cardiac disease mouse models and human cardiomyopathy, and the removal of *Runx1* in myocardial infarction attenuates adverse cardiac remodeling [220]–[223]. These studies provide validation that we have identified transcription factors that may be direct regulators of cardiac pathology and play an important role in the process. We focused our endeavors on examining the regulatory landscape for two shared transcription factors, *Klf15* and *Sox9*, that are shared in atrial fibrillation and heart failure, but the remaining candidate transcription factors may provide meaningful insight in cardiac pathophysiology. Some of the other transcription factors that are shared in the two mouse models and have been examined in a cardiac context include *Etv2*, *Foxm1*, *Lbh*, *Nupr1*, *Snai1*, and *Tcf21*. A candidate of interest that has not been examined in a cardiac context is *Foxs1* (Forkhead Box S1). Its functional role has been assessed in several cancer studies, where it appears to have an important role in Epithelial-to-Mesenchymal Transition (EMT) [224]–[227]. Foxs1 is highly expressed in the fibroblast and myofibroblast cells in the heart based on single-cell expression in Tabula Muris [228]. We hypothesize it may have an important role in cardiac fibroblasts, potentially with the activation of fibroblasts from the epithelium during cardiac remodeling. Our analysis suggests Foxs1 warrants further investigations into its role during cardiac disease.

Variations in non-coding regions is associated with disease, which implies defects in transcriptional regulation as a driver of disease risk. Based on the knowledge that noncoding RNAs (ncRNAs) are transcribed from regulatory elements, we have previously shown that the differential deep sequencing of ncRNAs from the wild-type and *Tbx5* mutant mice identified enhancers for calcium-handling genes, revealing a physiologically relevant gene regulatory network. Therefore, we performed ncRNA-seq profiling in the left atria of the pressure overload mouse model, and compared the two datasets. Previous analysis from our laboratory identified a differential ncRNA

transcript near *Ryr2*, which was shown to be *Tbx5*-dependent and necessary for gene expression. Our analysis has since confirmed that this ncRNA transcript and the *Ryr2* gene are also differentially expressed in the TAC mouse model. This suggests that the same enhancer is utilized to regulate the expression of *Ryr2* in both cardiac disease models. A more global analysis of the differential ncRNA transcripts had a positive correlation between the atrial fibrillation and heart failure mouse models, which suggests that we have identified a network of shared enhancers utilized in various cardiac disease contexts. The shared regulatory networks supports the notion that we are uncovering a generalizable disease-specific gene regulatory network that mediates the physiologic consequences of cardiac diseases.

The initial efforts in our laboratory provided important information regarding the regulatory mechanisms of cardiac ion channels and structural genes, but we were interested in examining regulatory mechanism for other cardiac disease processes. A ncRNA downstream of *Klf15* had decreased expression in both the *Tbx5*-deletion and TAC mouse models. *Klf15* has previously been shown to repress the expression of hypertrophy genes in cardiomyocytes, implicating it as an important transcriptional regulator in homeostasis [207], [229], [230]. *Klf15* is a negative regulator of cardiac hypertrophy through the inhibition of GATA4, MEF2, and myocardin binding to target genes, including ANF/BNP known markers of hypertrophy [207], [229], [230]. Non-coding transcripts can be used as a marker for active enhancer activity, and we propose the differential ncRNA downstream of *Klf15* is an enhancer-borne RNA. The *Klf15* enhancer aligned with ATAC-seq peaks, and other canonical marker for enhancer activity, including H3K27ac, H3K4me1, PolII, and p300. While the enhancer had all the appropriate markers for an active enhancer, we validated the enhancer was specifically regulating *Klf15* through the looping interactions in a micro-C dataset from cardiomyocytes. An interesting finding

from exploring this candidate enhancer was the binding of TBX5 at both the promoter and enhancer for *Klf15*. This novel finding suggests that TBX5 drives the expression of repressors that inhibit hypertrophy to maintain homeostasis in cardiomyocytes. This implies that the role of TBX5 is much greater than driving the expression of ion channels and gap junction proteins for cardiac rhythm, but rather it prevents cardiac disease pathways from being activated in homeostasis. This discovery leads us to speculate as to whether increased expression of TBX5 can prevent the pathophysiology observed in various cardiac perturbations. Another fascinating finding from our study was determining a human heart failure CAGE-seq peak lifted over to our *Klf15* enhancer. The implication that we have identified a evolutionarily conserved regulatory mechanism for *Klf15* suggests the importance of this enhancer/ncRNA, and warrants further investigation into the transcript in humans.

Our analysis also identified up-regulated ncRNAs that are activated in cardiac disease contexts, and are conserved in both the TAC and *Tbx5* KO mouse models. We suggest that their associated regulatory elements may reveal disease-response enhancers, essential for coping with atrial dysfunction. These candidate regulatory elements included a putative enhancer upstream of *Sox9*, a known modulator of cardiac fibrosis [211], [212]. The removal of *Sox9* ameliorated the fibrotic scarring observed from myocardial infarction and ischemic injury, implicating it as an important driver of fibrosis in cardiac disease [211], [212]. We propose the candidate ncRNA in the *Sox9* locus is also activated in the fibroblasts. The fibroblast ATAC-seq dataset had an open chromatin peak at the 5' end of the ncRNA. While there was ncRNA transcription, this putative enhancer did not have any other canonical markers for an active enhancer. We propose that this enhancer is activated in a disease-specific context only, so the fact that these active enhancer marks are not present is not surprising because the ENCODE datasets are generated in a wild-type

context. Performing H3K27ac or H3K4me1 ChIP-seq experiments in a cardiac disease mouse model may provide more meaningful markers for the enhancers and ncRNAs that are activated in a disease setting. To evaluate our hypothesis that the ncRNA is a disease-specific transcript in fibroblasts, we isolated cardiac fibroblasts and treated the cells with TGF-β, a canonical treatment to activate fibroblasts in an *in vitro* setting. The ncRNA-seq demonstrated our candidate was only expressed in the TGF-β treated fibroblasts, supporting the notion that the transcript is a disease-specific ncRNA. Examination of the ATAC-seq peaks in this locus reveals an opening chromatin peak at the 5' end of our candidate (Peak4). This open chromatin peak has the most activation in a luciferase reporter assay compared to the other cloned peaks in the locus, with increased activation following TGF-β treatment. All in all, our studies show we have identified a disease-specific ncRNA that is activated in both atrial fibrillation, heart failure, and activated fibroblasts. While we speculate the enhancer associated with our ncRNA is regulating the expression of *Sox9*, further evidence needs to be provided to connect our candidate enhancer to our gene of interest. If our enhancer is regulating the expression of *Sox9*, it is important to note that *Sox9* has been implicated in fibrotic processes in other tissues [213], [231], [232]; it would be interesting to determine whether this ncRNA is present in fibrosis of these other tissues or whether it is specific to the heart. Interestingly, the enhancer was tested in 3T3 mouse embryonic fibroblasts, implicating that this enhancer may be a general fibroblasts enhancer, that is not specific to the heart. While our interests have been in evaluating our candidate enhancer and its putative transcriptional regulators, it is important to note that we have not evaluated the functional role of this ncRNA transcript. As discussed in the introduction, many lncRNAs have been shown to be important in cardiac fibrosis [167], [168]. This is a very interesting avenue of research, as this transcript is activated in all of our cardiac perturbations, and future endeavors will be made to

provide clarity on the role of this ncRNA.

In summary, our studies have identified conserved transcriptional changes in atrial fibrillation and heart failure, and also shared regulatory mechanisms driving these changes. While our laboratory has made significant findings in the *Tbx5*-dependent gene regulatory network important for cardiac rhythm, it is clear that the shared mechanisms with heart failure play an instrumental role in the heart's ability to cope with atrial dysfunction. It is no longer simple to say "atrial fibrillation leads to heart failure and heart failure leads to atrial fibrillation," but rather there are connecting gene regulatory mechanisms that more closely align in these cardiac perturbations. For example, it is thought that atrial fibrillation that occurs in patients with heart failure is due to the cardiac remodeling and increased fibrosis that leads to abnormal conduction and atrial fibrillation. Our studies suggest that it may be more complicated than that, as there is a shared disease-specific gene regulatory network in our mouse models for atrial fibrillation and heart failure. This conservation of the transcriptional analysis and ncRNA candidates supports the paradigm of a common disease-specific gene regulatory network that mediates and may participate in the physiologic consequences of disease.

Chapter 3.5 Acknowledgments

The sham and transverse aortic constriction (TAC) surgeries were performed by David Park laboratory members (The Leon H. Charney Division of Cardiology, New York University Grossman School of Medicine). The HL-1 micro-C dataset was generated by the Ruthenburg lab (University of Chicago). The bioinformatic analysis was performed in conjunction with Zhezhen Wang (Moskowitz Lab, University of Chicago).

CHAPTER 4: TIME COURSE REMOVAL OF TBX5

Chapter 4.1 Introduction

The heart is composed of an assortment of cell types that function in sync with each other during homeostasis for proper cardiac function, but communication among these cell populations is equally as important during cardiac pathology [233]–[236]. While many signaling pathways have been implicated in the progression of heart failure, there is less understanding in how cardiac rhythm abnormalities lead to cardiac remodeling and fibrosis. Our studies aim to identify the earliest molecular changes in a time course deletion of *Tbx5*, our mouse model for atrial fibrillation. A time course RNA-seq of *Tbx5*-deficient atria will allow us to examine the progressive gene expression changes that occur during the transition of normal cardiac rhythm to spontaneous, sustained atrial fibrillation.

Our laboratory has previously shown the deletion of *Tbx5* leads to spontaneous atrial fibrillation [118]. Prior to the onset of atrial fibrillation, the examination of the transcriptional changes indicate the downregulation of ion channels confirming the importance of *Tbx5* in the regulation of calcium handling genes [118]. Even though our initial RNA-seq was collected prior to the onset of atrial fibrillation, we have already identified differential gene expression indicative of transcriptional changes in other cell types. Gene ontology analysis of the upregulated genes revealed an enrichment for terms related to changes in the fibroblasts (Extracellular matrix organization GO:0030198) and immune cells (inflammatory response GO:0006954). Since the deletion of *Tbx5* is specific to cardiomyocytes, we hypothesize there may be a paracrine signal from the cardiomyocyte to the other cell types in the heart.

The initial RNA-seq had transcriptional changes indicative of disease processes in other cell types, and we believe this may be too far down the timeline from the initial perturbation.

Therefore, we generated a time course deletion of *Tbx5* where tissue was collected at 4 time points (Day 3, Day 6, Day 10 and Day 17) following the removal of *Tbx5* with the start of the tamoxifen regiment. We performed RNA-seq experiments in the left atria at each of these time points, and our initial efforts were spent on significant gene expression changes between the control and *Tbx5*-deficient left atria through the time course. The Day 3, Day 6, and Day 10 time points are prior to the onset of atrial fibrillation, while the Day 16 time point is around the time period when the mice are in spontaneous, sustained atrial fibrillation. An interesting trend demonstrated that the removal of *Tbx5* resulted in an early transcriptional response at Day 3 and Day 6, which is very distinct from the transcriptional response at Day 10 and Day 17. The later timepoints reveal transcriptional changes indicative of a disease network activation, which includes extracellular matrix remodeling and an immune response. This is in contrast to the transcriptional changes at the early time points, which were related to transcriptional changes specific to the cardiomyocytes. maSigPro analysis provided us with promising candidate gene targets, which will be evaluated by our laboratory in future studies. For example, *Hopx* is downregulated in the time course starting at Day 6, and Lbh is upregulated beginning at Day 3. Our analysis has also identified mis-regulated ligands at the various time points, which has given us a list of candidate signaling pathways involved in the progression of atrial fibrillation. Future experiments, will examine whether the modulation of these candidates will prevent the pathophysiology that occurs in atrial fibrillation following the deletion of *Tbx5*.

Chapter 4.2 Materials and Methods

Control and Tbx5-mutant coding RNA-seq library preparation and sequencing

Generation of the conditional *Tbx5* deletion was completed with a floxed allele, and was previously described by our laboratory [64], [118], [159]. The third exon of *Tbx5* was excised with a tamoxifen inducible Cre recombinase driven by the R26 locus. $Tbx5^{fl/fl}$;$R26^{CreERT2}$ and $R26^{CreERT2}$ mice were subjected to a tamoxifen regiment over 3 days at 6-10 weeks of age, as previously described [118]. The mice used in these experiments were raised in accordance with the Guide for the Care and Use of Laboratory Animals. All experiments were approved by the University of Chicago Institutional Animal Care and Use Committee (IACUC).

Left atrial tissue from $R26^{CreERT2}$ (control) and $Tbx5^{fl/fl}$;$R26^{CreERT2}$ (*Tbx5*-mutant) mice was dissected at 4 time points after the initiation of the tamoxifen injections: Day 3, Day 6, Day 10, Day 17. The tissue was mechanically homogenized in TRIzol Reagent. Total RNA was prepared with a TRIzol-based extraction and isolated using a Qiagen RNeasy Mini Column as previously described [118], [159]. PolyA selection was performed using the Lexogen CORALL PolyA selection module, and subsequently library preparation was completed with the Lexogen CORALL Total RNA Library Kit. The libraries were sequenced on the Illumina Hi-Seq machinery. Samples were sequenced 50bp single-ended at 10-15 million reads per replicate. The reads were mapped to mm10 using STAR v2.5.3(UCSC mm10 annotation). Genomic read alignments (BAM files) were filtered by MAPQ score greater than 30 with samtools v1.9, and counts were generated for each sample with HTSeq-counts (v.0.11.2) with mm10 known gene annotation from UCSC [196]. Raw counts were normalized by Relative Log Expression (RLE) in DESeq2, and genes with counts larger than 10 in at least 3 samples are kept for downstream analysis. Then, raw counts of the expressed genes are used by the default DESeq2 analysis pipeline to identify differentially

expressed genes between control and *Tbx5*-mutant mice [197]. Genes with a |log2 fold change| greater than 0.5, and adjusted p-value less than 0.05 are considered significantly differentially expressed.

Gene ontology enrichment analysis

Differential gene lists are used as input for gene ontology analysis using Metascape (https://metascape.org).

maSigPro Time Course Analysis

maSigPro is a regression based approach to identify these significant gene expression changes in a time course RNA-seq. maSigPro analysis was performed to identify significant gene expression changes between control ($R26^{CreERT2}$) and *Tbx5*-mutant ($Tbx5^{fl/fl}$;$R26^{CreERT2}$) samples at Day 3, Day 6, Day 10, and Day 17 after initial dose of tamoxifen [237], [238]. In this cluster analysis, the elbow plot was used to identify the appropriate number of clusters.

Heatmap of time-course Tbx5-deletion RNA-seq

Differentially expressed genes are pooled together to create a gene list. The log2 Fold Change of those genes are plotted into a heatmap with non-squared Euclidean distances and Ward2 hierarchical clustering method. The heatmap is than divided into 16 groups based on the clustering.

Ligand Identification

Ligands were identified using CellTalkDB database, a curated list of ligand-receptor interactions in humans and mice [239]. Our differential gene lists were overlapped with the annotated ligands from this database.

Chapter 4.3 Results

Time-course deletion of Tbx5 implicates an early transcriptional response that is distinct from the later response

Our laboratory has previously demonstrated the conditional deletion of *Tbx5* leads to spontaneous, sustained atrial fibrillation. To identify the earliest transcriptional changes that occur in the atrial fibrillation mouse model, we performed transcriptional profiling of the left atria from *Tbx5*-deficient and control mice in a time course RNA-seq. The left atria from *Tbx5*-deletion ($Tbx5^{fl/fl}$; $R26^{CreERT2}$) and control ($R26^{CreERT2}$) mice were collected at Day 3 (red), Day 6 (green), Day 10 (blue), and Day 17 (purple) following the initiation of the tamoxifen regiment (Figure 4.1A). At the Day 3 and Day 6 timepoints, a comparison of the control and *Tbx5*-deficient atria showed fairly discrete expression patterns between the two genotypes, and fairly consistent expression among the biological replicates (Figure 4.1B-C). To visualize the variance among the replicates, principal component analysis (PCA) for the Day 3 and Day 6 samples revealed a separation on principal component 1 (PC1) due to genotype (Figure 4.1B-C). PC1 accounted for 55% and 56% of the variance among the replicates in Day 3 and Day 6, respectively (Figure 4.1B-C). It is interesting to note the separation between the genotypes is much larger at Day 6, implicating more distinct variation between the samples. Heatmaps of the Day 10 and Day 17 timepoints reveal consistent expression patterns amongst the biological replicates of the control and *Tbx5*-deficient samples (Figure 4.1D-E). The PCA plot for the Day 10 and Day 17 time points reveal >80% of the variance is accounted for by genotype on PC1 (Figure 4.1D-E). All the samples for each time point were shown in a single PCA plot (Figure 4.1F). The *Tbx5*-deficient samples at Day 10 and Day 17 separated from the remainder of the samples on PC1, while PC2 separates the Day 3 and Day 6 samples based on genotype. This suggests the transcriptional changes that occur

at Day 10 and Day 17 are similar to each other and distinct from the transcriptional changes at the earlier timepoints. The transcriptional profiling has therefore identified an early and late response that are very different from each other in the atrial fibrillation model.

Gene ontology analysis reveals unique differential gene expression in the early and late time points

While the PCA plot suggests there is a transcriptional difference between the early and late timepoints, it was necessary to examine the differentially expressed genes at each time point independently. The volcano plot for the Day 3 timepoint visualizes the differentially expressed genes, and shows 629 genes are significantly changed with the removal of *Tbx5*. The red dots represent the 374 significant down-regulated genes (log2FoldChange<0.5, p.adj<0.05), the blue dots are represent the 255 significant up-regulated genes (log2FoldChange>0.5, p.adj<0.05), and the grey dots did not reach the threshold for significance (Figure 4.2B). Gene ontology analysis of the down-regulated genes resulted in an enrichment of terms related to disease process, including "negative regulation of immune system process (GO:0002683)", "wound healing (GO:0042060)", and "cytokine production (GO:0001816)" (Figure 4.2A). This is in contrast to the up-regulated genes enriched for gene ontology terms related to metabolism and conduction ("Fatty acid metabolism (R-MMU-8978868)"; "Muscle contraction (R-MMU-397014)"; Figure 4.2C). Even at the earliest time point we can see that the removal of *Tbx5* results in gene expression changes that can affect the functionality of the heart. The Day 6 time point showed 1,107 differentially expressed genes after the removal of *Tbx5* (559 genes log2FoldChange<0.5, p.adj<0.05; 548 genes log2FoldChange>0.5, p.adj<0.05; Figure 4.2E). The gene ontology analysis for the up-regulated genes at this time point revealed very similar differential gene expression to the Day 3 time point; there was an enrichment for terms related to metabolic processes and cardiac conduction. Interestingly, when we examine the down-regulated genes at the Day 6 time point they are

predominately enriched for gene ontology terms related to cardiac conduction, and we no longer identified gene ontology terms indicative of disease processes. The Day 10 timepoint resulted in 2,877 differentially expressed genes as shown in the volcano plot (1410 genes, log2FoldChange<0.5, p.adj<0.05; 1467 genes, log2FoldChange>0.5, p.adj<0.05; Figure 4.2H). The analysis of the Day 17 timepoint revealed 2,215 genes are significantly changed after the removal of *Tbx5* (1079 genes, log2FoldChange<0.5, p.adj<0.05; 1136 genes, log2FoldChange>0.5, p.adj<0.05; Figure 4.2K). The gene ontology analysis of the differentially expressed genes at these two time points were very interesting because they were enriched for very similar terms. The down-regulated genes were enriched for terms related to cardiac conduction and contraction (Figure 4.2G,J; "regulation of ion transport GO:0043269", "heart contraction GO:0060047", "Muscle contraction R-MMU-397014"). TBX5 is directly regulating the expression of these genes, so in the absence of *Tbx5* we would expect to see an enrichment of these cardiac rhythm terms in the analysis for the down-regulated genes. Interestingly, the gene ontology analysis for the up-regulated genes was enriched for terms related to extracellular matrix and collagen formation (Figure 4.2I,L; "Extracellular matrix organization R-MMU-1474244", "Collagen formation R-MMU-1474290"). These terms are predominantly related to fibroblast activity, and in many disease settings there are injury repair processes that involve the activation of the fibroblasts that can produce an excessive amount of extracellular matrix proteins. In our mouse model spontaneous, sustained atrial fibrillation is present at about Day 17 and beyond, so it is fascinating that genes related to fibrotic disease processes are appearing at Day 10, prior to the onset of atrial fibrillation or the fibrotic remodeling .

Identification of differential gene expression patterns through the Tbx5 deletion time course

We next interrogated the changes in expression patterns by plotting the log2 fold change of the differentially expressed genes using a hierarchical clustering method. The differentially expressed genes were separated into 16 clusters and gene ontology analysis was evaluated for the genes in each of the clusters. A closer inspection of the patterns reveals the peach and yellow clusters include genes that are down-regulated in the earliest time point, and continue to be down-regulated throughout the time course. The clusters are enriched for gene ontology terms related to cardiac conduction, therefore confirming that the earliest transcriptional changes are due to the changes in the genes that control atrial rhythm (Figure 4.S1, "cardiac conduction GO:0061337" and "positive regulation of potassium ion transmembrane transport GO:1901381"). To further explore the early transcriptional changes, we examined the teal pattern that is up-regulated in the earliest time point, and continues to be upregulated throughout the time course. The "regulation of interleukin-6 production" is particularly interesting as this pathway has been implicated in disease. IL-6 is a proinflammatory cytokine and chronic, long-term IL-6 signaling can lead to fibrosis and inflammation [240]. Lastly, when we examine the purple clusters with increased gene expression at the later time points (Day 10 and Day 17), it is enriched for terms related to fibroblast activity (Figure 4.S1; "Collagen biosynthesis/modifying enzymes R-MMU-1650814"; supramolecular fiber organization GO:0097435"). This analysis highlights the differences that are occurring through the time course RNA-seq, and demonstrates a large shift in the transcriptional changes that begin to include other cell types of the heart around Day 10.

Identification of candidate genes that are significantly changing throughout the RNA-seq time course

To further characterize the changes that occurred through the time course, we used maSigPro to identify candidate genes that are significantly changed in this time course RNA-seq.

maSigPro is a regression based approach that identifies genes with significant expression changes between the genotypes and also temporally [237]. The analysis separates the significant genes into 4 clusters shown in Figure 4.3. Cluster 1 identifies genes with significant decreased expression throughout the time course, and gene ontology analysis shows an enrichment for "Muscle contraction (R-MMU-397014)" and "cardiac conduction (GO:0061337)". Previous analysis has also shown this enrichment, but maSigPro identifies gene candidates that follow this trend. Cluster 1 has 816 genes, and include genes for calcium handling, ion channels, and gap junctions including *Ryr2*, *Atp2a2*, *Sln*, *Scn5a*, and *Gja5*. Many of these genes have been previously studied by our lab and shown to be direct targets of TBX5. In addition to cardiac conduction genes in Cluster 1, we have identified other candidates that could play an important role in the pathophysiology of atrial fibrillation. *Klf15* is a significantly differential gene between the *Tbx5*-deletion and control mice that was discussed in Chapter 3. This analysis now shows that *Klf15* also has significant gene expression changes temporally. Another candidate gene from Cluster 1 that is of particular interest to our laboratory is *Hopx*. The role of *Hopx* has previously been studied in the developing mouse heart, and *Hopx*-deficient mice had poorly developed myocardium in the embryo or reduced cardiac function in the adult, with variable penetrance of disease.

Cluster 3 and Cluster 4 identify candidate genes with increased gene expression through the time course (Figure 4.3C-D). Cluster 3 contains 487 genes, including Ankrd1, Foxm1, and Tbx18. Cluster 4 contains 554 genes, including Lbh, Meox1, Nupr1, and Sox9. Gene ontology analysis for both clusters showed an enrichment for terms very similar to those seen in the individual analysis of the upregulated genes in the Day 10 and Day 17 timepoints; the terms are related to fibroblast function, including "supramolecular fiber organization (GO:0097435)" and "Extracellular matrix organization R-MMU-1474244" (Figure 4.3C-D). A literature search has

shown that many of the genes identified in this analysis have been previously implicated in disease processes. For example, *Meox1* has recently been shown to be an essential transcriptional activator in the transition of fibroblasts to myofibroblasts. *Lbh* is a transcription factor important during embryogenesis for heart and limb development, but the role of this transcription factor in the adult mouse during heart disease needs to be further elucidated. In our time course, Lbh is differentially expressed at the Day 3 timepoint, and continues to be significantly up-regulated at each subsequent timepoint (Day 3 log2FoldChange=0.55, adj. p-value=0.032; Day 6 log2FoldChange=1.01, adj. p-value=1.65E-9; Day 10 log2FoldChange=1.71, adj. p-value=2.87E-22; Day 17 log2FoldChange=1.88, adj. p-value=4.21E-27). Lbh is only 1 of 3 transcription factors that is up-regulated from the earliest time point and continues that trend throughout the time course. In summary, the maSigPro analysis identified gene candidates that may be crucial in driving the pathophysiology of atrial fibrillation, and future work will be directed at exploring their role in cardiac disease.

Identification of the differentially expressed ligands in the atrial fibrillation mouse model

Gene ontology analysis of the early timepoints in the time course suggest that the transcriptional changes are confined to the cardiomyocytes, which makes sense because *Tbx5* is expressed in the cardiomyocytes. The later timepoints in the time course implicate transcriptional changes in other cell types including fibroblasts and immune cells, implicating a cell non-autonomous interaction. For our time course deletion of *Tbx5* we know the initial perturbation is in the cardiomyocyte, therefore we next wanted to explore the ligands that may be responsible for triggering the transcriptional changes in the other cell types of the heart. Therefore, we identified the top 15 down-regulated and up-regulated ligands at each of the respective time points. A closer examination of the candidates reveals Lgals3 and Bmp3 are two ligands which are down-regulated

at Day3 through Day 17. Lgals3 (Gal-3) is a B-galactoside-binding lectin that has been shown to be a putative marker for cardiac fibrosis and inflammation. On the other hand, *Bmp3* has not been examined in the context of cardiac disease. *Bmp3* is down-regulated at D3, and it is the most down-regulated ligand in the subsequent timepoints (Day 3 log2FoldChange=-0.96, adj. p-value=0.00018; Day 6 log2FoldChange=-3.8, adj. p-value=7.9E-31; Day 10 log2FoldChange=-5.37, adj. p-value=5.72E-29; Day 17 log2FoldChange=-6.64, adj. p-value=1.98E-35). The degree of Bmp3 mis-regulation at each time point in the RNA-seq time course makes it an interesting candidate, but it will require future investigations to determine the requirement of this ligand. As we continue to search for a ligand that activates the downstream atrial fibrillation pathophysiology and the associated cardiac remodeling, our analysis has identified candidates that may contribute to this process.

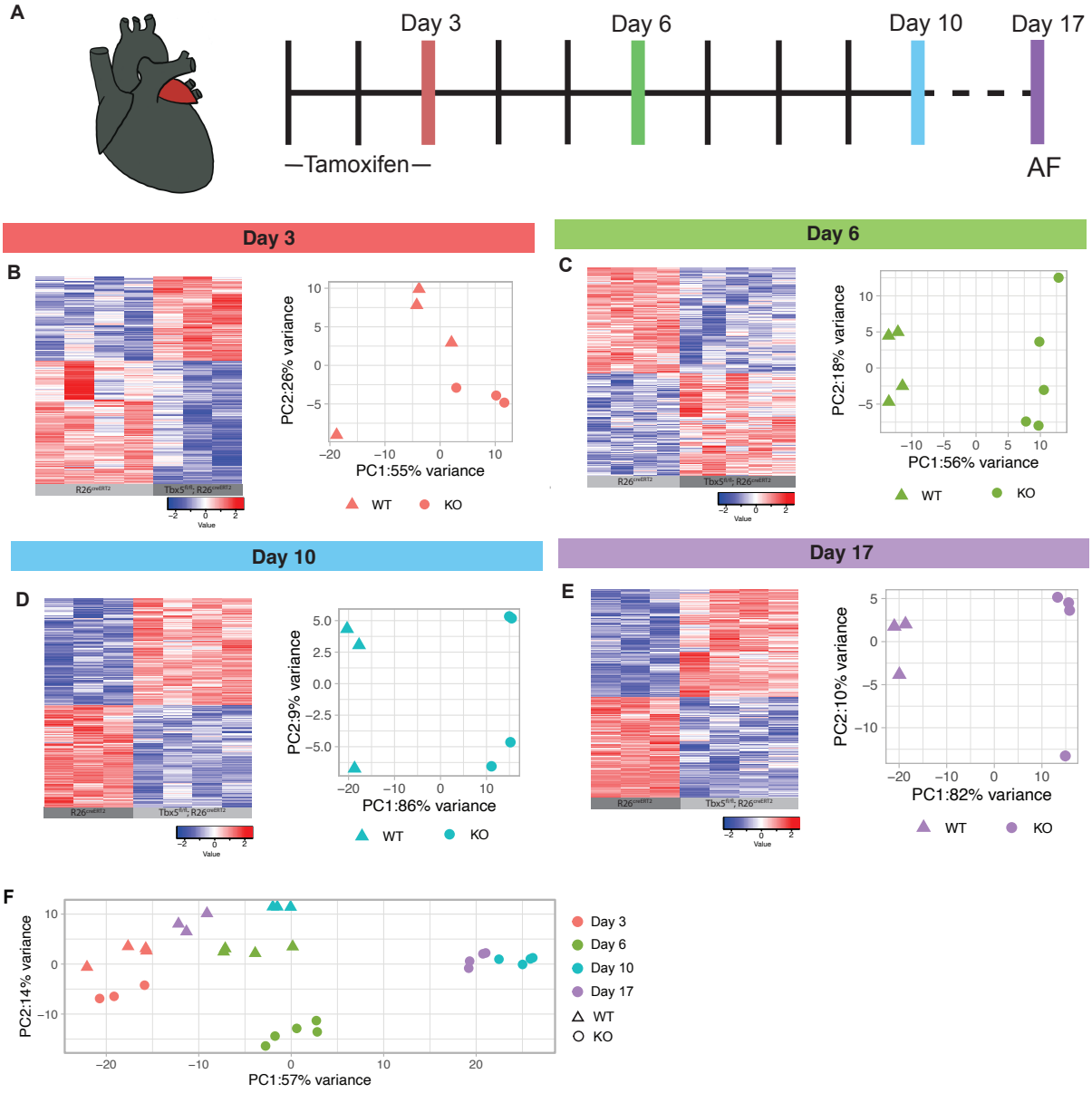


Figure 4.1. Deletion of TBX5 from different time points suggests the presence of an early and late response. A. Schematic highlighting the left atria was used for RNA-seq (left). Timeline demonstrating 3 days of tamoxifen treatment, followed by the collection of the atria at Day 3, Day 6, Day 10, and Day 17. B-E. Left atria collected for RNA-seq at Day 3 (B), Day 6 (C), Day 10 (D), and Day 17 (E). Heatmap for each respective time point of the significantly mis-regulated transcripts in the left atria of control ($R26^{CreERT2}$) vs. TBX5 KO ($Tbx5^{fl/fl};R26^{CreERT2}$) (left). Principal component analysis (PCA) plot for each respective time point (right). F. PCA plot for all the samples from the time course RNA-seq. Day 10 and Day 17 replicates are separated by genotype on PC1, while Day 3 and Day 6 are separated by genotype on PC2.

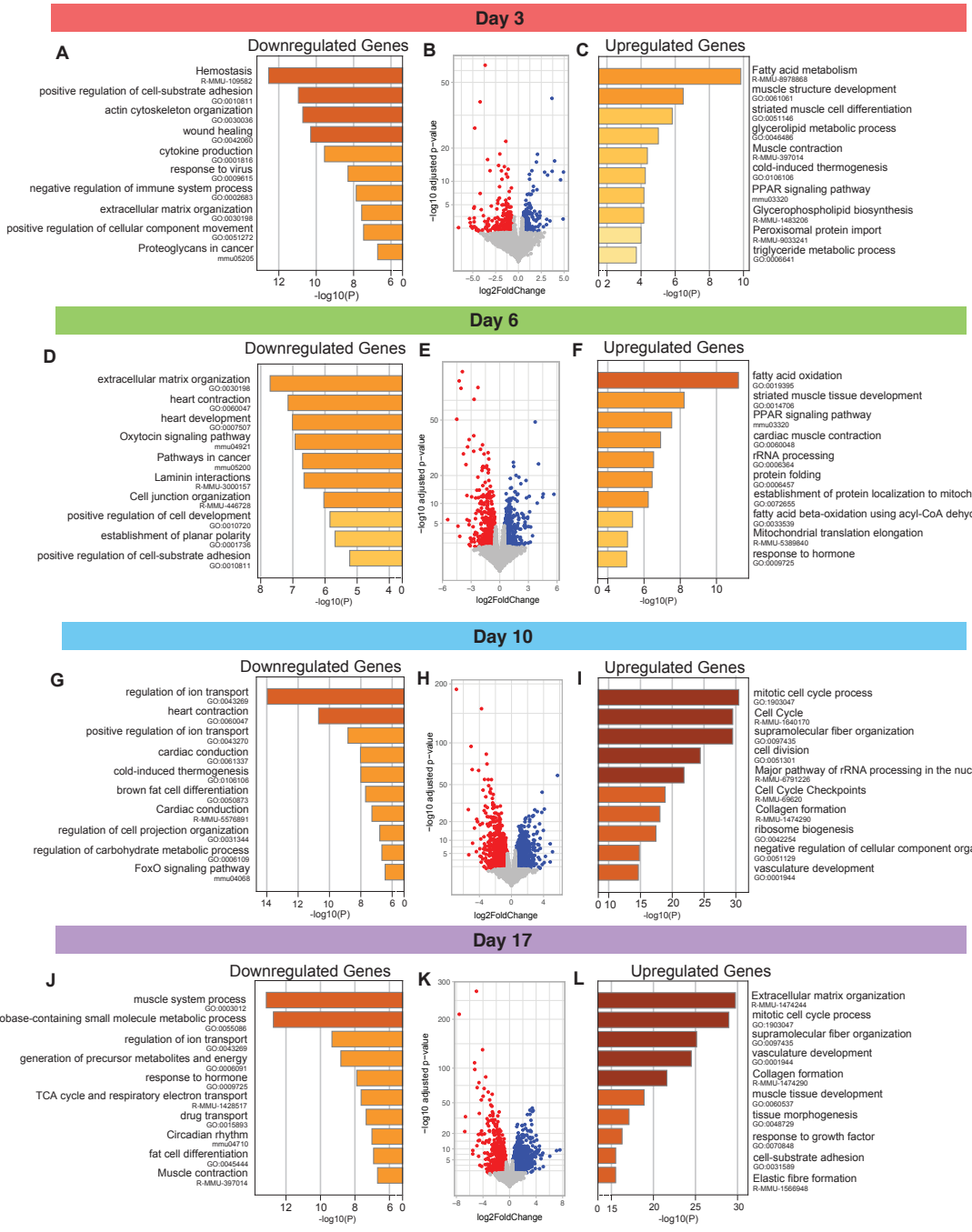
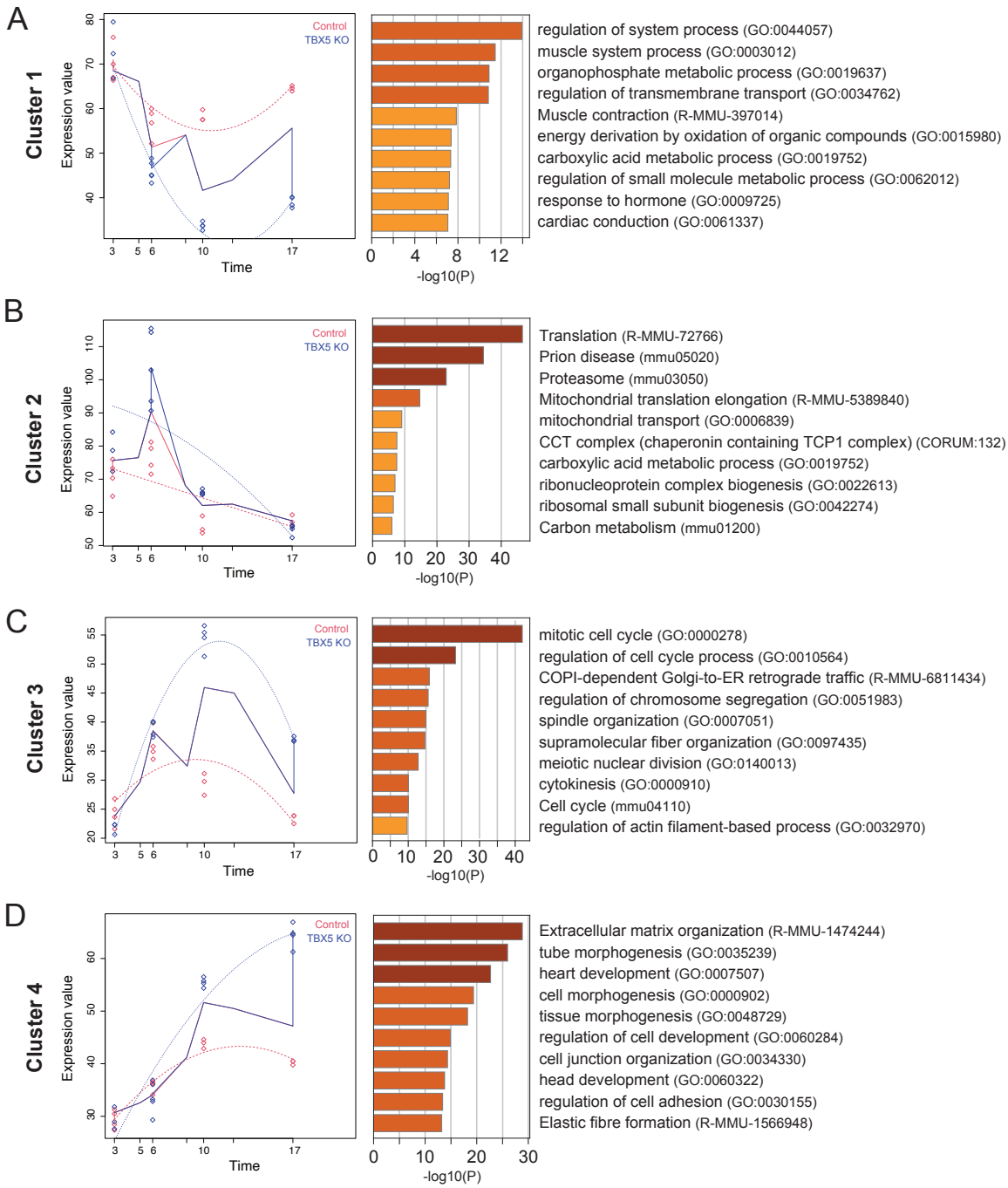


Figure 4.2. Differential gene expression reveals different disease mechanisms in the early time points compared to the later time points in the TBX5 deletion time course. A-L. Differential gene expression comparing control (*R26*^{CreERT2}) vs. TBX5 KO (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) was examined for each of the time points: Day 3 (A-C), Day 6 (D-F), Day 10 (G-I), Day 17 (J-L). Gene ontology analysis on significant down-regulated genes at Day 3 (A), Day 6 (D), Day 10 (G), Day 17 (J). Volcano plot (log2FoldChange vs. -log10(adj. p-value)) of control (*R26*^{CreERT2}) vs. *Tbx5* KO (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) differentially expressed (DE) transcripts in RNA-seq for Day 3 (B, 629 DE

Figure 4.2, continued

genes), Day 6 (E, 1107 DE genes), Day 10 (H, 2877 DE genes), Day 17 (K, 2215 DE genes). Significantly downregulated genes (log2FoldChange<0.5 and adj. p-value<0.05) are shown in red. Significantly upregulated genes (log2FoldChange>0.5 and adj. p-value<0.05) are shown in blue. Non-significant differentially expressed transcripts are grey. Gene ontology analysis on significant up-regulated genes at Day 3 (C), Day 6 (F), Day 10 (I), Day 17 (L).



E

Cluster 1	816 genes	Ryr2, Atp2a2, Gja5, Scn5a, Sln, Kcnj3, Etv1, Hopx, Fgf12
Cluster 2	355 genes	Rps2, Slc25a4, Psmb8, Csnk2b,Socs3
Cluster 3	487 genes	Ankrd1, Basp1, Creb3, Foxm1, Irf7 Tbx18, Tcf19, Zfpm1
Cluster 4	554 genes	Lbh, Meox1, Nupr1, Pknox2, Rest, Runx1, Sox9, Tcf21, Tcf4, Twist1

Figure 4.3. Identification of significant differential gene expression changes through the TBX5 deletion time points. A-D. maSigPro analysis for RNA-seq comparing control (*R26*^{CreERT2}) vs.

Figure 4.3, continued

TBX5 KO (*Tbx5*$^{fl/fl}$;*R26*^{CreERT2}) at 4 time points: Day 3, Day 6, Day 10, and Day 17 after tamoxifen regiment initiated. (A) Cluster 1 includes genes (n=816 genes) that are down-regulated throughout the time course, and gene ontology analysis sorted by -log10(p-value). (B) Cluster 2 includes genes (n=355 genes) changed in the early time points, and gene ontology analysis sorted by -log10(p-value). (C) Cluster 3 includes genes (n=487 genes) that are up-regulated throughout the time course, and gene ontology analysis sorted by -log10(p-value). (D) Cluster 4 includes genes (n=554 genes) that are up-regulated throughout the time course, and gene ontology analysis sorted by -log10(p-value). E. Table displaying the cluster, the number of genes identified in each cluster, and candidate genes of interest identified in the cluster.

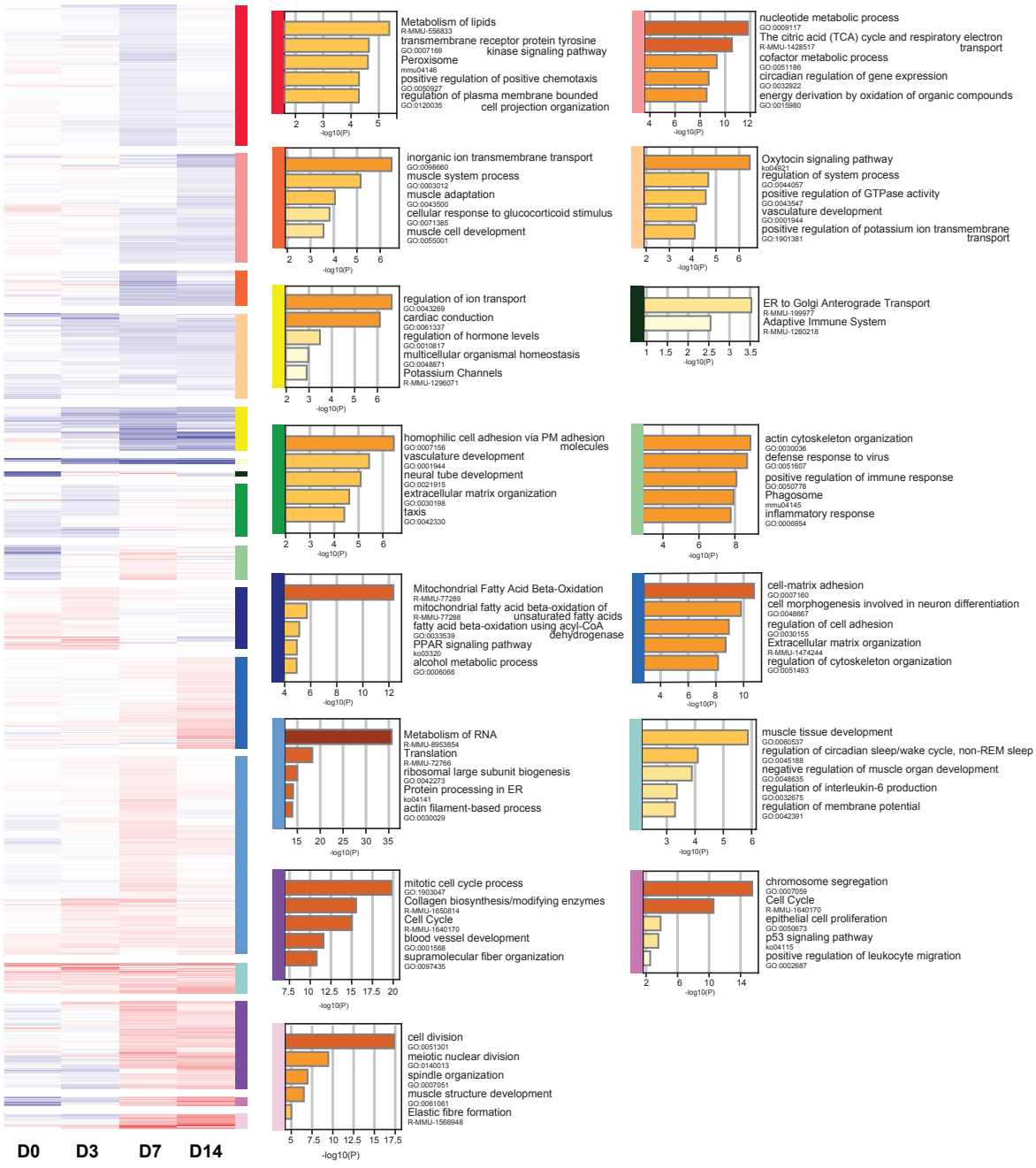


Figure 4.S1. Identification of differential gene expression patterns through the TBX5 KO time points. Heatmap with hierarchical clustering of differential gene expression for the TBX5 deletion timecourse. Each cluster has a color that corresponds to gene ontology analysis.

A

Gene	Day 3 log2FC	adj. p-value
Lgals3	-4.19	2.071E-40
Nts	-4.18	6.455E-08
Thbs1	-2.96	1.679E-05
Prg4	-2.81	0.0187964
Selplg	-1.93	0.0282349
Cx3cl1	-1.62	0.0003928
Ang	-1.50	0.0108856
Ptpn6	-1.18	0.0489088
Alox5ap	-1.18	0.0226761
Pdgfb	-1.01	0.0191988
Bmp3	-0.96	0.0001801
Lgals3bp	-0.91	0.0015815
Vasp	-0.91	0.0013843
Sfrp1	-0.87	0.0008227
Tnfsf10	-0.81	0.039198

Gene	Day 3 log2FC	adj. p-value
Il6	2.7442843	0.0135907
Fgf10	2.2385262	0.0323886
Bmp5	2.085768	0.0015997
Igf2	1.6104935	1.699E-08
Bdnf	1.5524226	0.0287035
Cxcl14	1.3766901	0.0235854
Egf	1.0427924	0.0011536
Sema3a	0.9951892	0.0124608
Il15	0.9020766	0.0015997
Fgf1	0.8524016	6.253E-05
Lama2	0.5576894	0.0004645

B

Gene	Day 6 log2FC	adj. p-value
Bmp3	-3.8038387	7.938E-31
Lgals3	-3.2074999	1.319E-38
Sema6c	-2.2051526	1.227E-08
Cdh1	-2.0414756	0.0308984
Tg	-1.9732691	2.151E-08
Nts	-1.6312777	0.0028356
Inha	-1.6132433	0.0001572
Sfrp2	-1.5980628	7.895E-11
Angpt1	-1.5405463	1.076E-06
Nppa	-1.3190594	3.765E-13
Ptn	-1.2799313	0.0026235
Sost	-1.1882883	0.0197389
Gng2	-1.138569	1.059E-18
Nrg2	-1.1363018	0.0040943
Mfap2	-1.0689536	0.0112293

Gene	Day 6 log2FC	adj. p-value
Angptl4	3.2429561	5.641E-05
Serpine1	2.7590333	1.059E-18
Rbp4	2.6654581	8.384E-05
Ccl7	2.6477624	1.774E-05
Ccl2	2.5582839	3.64E-05
Cxcl14	2.4731187	1.427E-07
Hbegf	1.832027	1.919E-11
Adam12	1.7482003	0.0298217
Lcn2	1.6239652	0.0213564
Tnc	1.6093288	0.0004696
Ereg	1.4724966	0.0024661
Bdnf	1.3553689	0.0137101
Inhbb	1.2554026	0.0031051
Timp1	1.2235143	7.557E-07
Fgf1	0.98379	3.37E-09

C

Gene	Day 10 log2FC	adj. p-value
Bmp3	-5.3730223	5.722E-29
Gnb3	-3.569159	0.0003806
Ccl19	-3.3748824	0.0211069
Mmp9	-3.1370653	5.816E-07
Cntf	-2.8683291	0.0012681
Lgals3	-2.8200326	8.977E-24
Dkk3	-2.332218	1.704E-32
Nrg2	-2.2560992	2.938E-11
Sema6c	-2.0476489	8.484E-09
Angpt1	-1.8966258	3.063E-08
Tg	-1.6710334	7.424E-06
Pla2g5	-1.6580452	7.775E-11
Vtn	-1.541412	5.145E-06
Fgf12	-1.4830792	5.717E-17
Nts	-1.4629724	0.0044285

Gene	Day 10 log2FC	adj. p-value
Tnc	5.7404807	3.075E-60
Thbs4	4.8332835	3.382E-09
Gdf6	3.9721446	1.405E-15
Serpine1	3.777857	6.16E-27
Cthrc1	3.579509	1.355E-11
Ereg	3.5056479	2.401E-19
Gdf10	3.3662595	0.0001521
Thbs1	2.958145	1.273E-06
Bdnf	2.833067	3.401E-09
Postn	2.7429962	7.902E-11
Edn1	2.7375462	2.007E-09
Il6	2.7007691	0.0010163
Timp1	2.6544944	2.529E-24
Bmp4	2.6113737	3.492E-24
Ccl7	2.4992729	0.0007677

D

Gene	Day 17 log2FC	adj. p-value
Bmp3	-6.638457	1.983E-35
Lgals3	-3.1377176	7.868E-27
Ccl11	-3.1203104	0.0026584
Mmp9	-3.1123171	8.162E-06
Cxcl1	-2.8625169	0.0071753
Gnb3	-2.4159469	0.0013957
Cntf	-2.4053668	0.0002784
Tg	-2.2842772	7.649E-09
Nrg2	-2.1658386	1.08E-10
Dkk3	-2.1634118	3.974E-28
Sema6c	-2.0275414	1.946E-09
Ahsg	-1.993034	5.055E-05
Fgf12	-1.9239576	4.398E-28
Gng2	-1.7579017	3.16E-36
Wnt11	-1.582785	0.0003815

Gene	Day 17 log2FC	adj. p-value
Lif	4.5252172	0.0063151
Gdf6	4.1090363	4.672E-16
Edn1	4.0060589	2.847E-19
Thbs4	3.8176791	0.0002304
Timp1	3.2272003	6.103E-32
Tnc	3.2053024	1.985E-14
Spon2	3.0831419	1.589E-22
Nrg1	3.082455	3.01E-09
Angptl4	3.0758639	0.0014961
Sfrp2	3.0617204	2.799E-42
Ereg	2.742559	1.79E-08
Cd24a	2.6143949	3.012E-19
Sost	2.6034341	1.302E-08
Col8a1	2.5916892	3.594E-09
Ada	2.5473206	1.477E-05

Figure 4.S2. Identification of the most highly mis-expressed ligands at each time point. A. Top 15 down-regulated ligands (left) and only 11 up-regulated ligands (right) at Day 3 following the removal of *Tbx5*, sorted by log2FoldChange. B. Top 15 down-regulated ligands (left) and top 15 up-regulated ligands (right) at Day 6 following the removal of *Tbx5*, sorted by log2FoldChange. C. Top 15 down-regulated ligands (left) and top 15 up-regulated ligands (right) at Day 10 following the removal of *Tbx5*, sorted by log2FoldChange. D. Top 15 down-regulated ligands (left) and top 15 up-regulated ligands (right) at Day 17 following the removal of *Tbx5*, sorted by log2FoldChange.

Chapter 4.4 Discussion

Transcriptional profiling of the left atria from *Tbx5*-deletion (*Tbx5*$^{fl/fl}$; *R26*^{CreERT2}) and control (*R26*^{CreERT2}) mice found TBX5 directly drives the expression of cardiac ion channels and structural genes. While this research was instrumental in understanding the functional role of TBX5 in maintaining cardiac rhythm homeostasis, it was evident there were also transcriptional changes in other cell types, including the fibroblasts. This is a fascinating observation that demonstrates the transcriptional changes related to fibrosis occurs much earlier than the onset of atrial fibrillation and well in advance of the phenotypic signs of fibrosis. This leads to an interesting question of whether the electrical conduction defects can be uncoupled from the fibrotic processes in cardiac disease. Even though our initial RNA-seq was collected prior to the onset of atrial fibrillation, we felt it was necessary to look earlier in the transition to atrial fibrillation to identify the earliest transcriptional and signaling changes. RNA-seq experiments were performed at Day 3, Day 6, Day 10, and Day 17 after the initiation of the tamoxifen regiment. Spontaneous, sustained atrial fibrillation appears around Day 17, therefore the Day 3, Day 6, and Day 10 timepoints are from tissues collected prior to the onset of atrial fibrillation. Interestingly, there seems to be a transition in the disease processes that occurs between Day 6 and Day 10 because we begin to see transcriptional changes indicative of fibroblast activity during this time period. Our analysis of the time course RNA-seq identified a transcriptional modulator and a signaling molecule that warrant future studies in for their role in the progression of cardiac pathology in atrial fibrillation.

A time course analysis using maSigPro identified transcriptional changes that are significant within each time point, and also significantly changed temporally. We speculate the list of candidate genes in this analysis identify important changes that occur during the transition from a healthy to atrial fibrillation heart, and potentially identifies the drivers of the pathological

processes. Cluster 1 is characterized by genes that are down-regulated through the timecourse, and many genes within this cluster are direct TBX5 target genes related to cardiac rhythm, including *Ryr2*, *Atp2a2*, and *Scn5a*. Cluster 1 also included *Hopx* (HOP Homeobox), a small homeodomain protein that does not bind to DNA. *Hopx* is highly expressed in the embryonic and adult heart, and the deletion of *Hopx* resulted in partial embryonic lethality [241], [242]. Their studies suggested that HOPX was a negative modulator of SRF (serum response factor) to modulate the expression of its downstream targets by recruiting HDAC proteins [241]–[243]. Subsequent studies demonstrated that the overexpression of *Hopx* resulted in cardiac hypertrophy, while *Hopx* null mice that survived the partial embryonic lethality did not display hypertrophy [243]. Based on our transcriptional profiling experiments, we would have expected the *Hopx* null mouse to have hypertrophy, therefore further investigations are warranted to examine the function of *Hopx* in the adult heart. For example, a cardiomyocyte-specific conditional deletion of *Hopx* from the adult mouse would provide an unbiased approach to determine the effect *Hopx* deficiency in the adult. Further elucidating the role of *Hopx* during cardiac perturbations is imperative as it has been shown to have decreased expression in human and mouse heart failure [244]. Our collaborators in the laboratory of Igor Efimov (Northwestern University) generated a human heart failure CAGE-seq dataset, and we also saw this change in the *Hopx* promoter. Therefore, our time course analysis has identified a candidate transcriptional modulator that has relevance to human cardiac disease. Evaluation of our timecourse RNA-seq reveals that *Hopx* is down-regulated at the D6, Day 10, and Day 17 timepoints (Day 6 log2FoldChange -0.85, adj. p-value 2.6E-6; Day 10 log2FoldChange -2.6, adj. p-value 1.3E-42; Day 17 log2FoldChange -3.2, adj. p-value 2.3E-64). *Hopx* is down-regulated in heart failure and atrial fibrillation, which means it may be involved in a shared transcriptional response for cardiac stress. The change in *Hopx* expression begins at the

Day 6 timepoint, suggesting that the change in expression is happening at an early timepoint and well in advance of the onset of atrial fibrillation. A closer expectation of the *Hopx* locus in our datasets also demonstrates that the *Hopx* promoter and a upstream enhancer have TBX5 ChIP-seq binding sites. Previous publications have shown that *Hopx* is regulated by Nkx2-5, which suggests *Hopx* is a target gene of canonical cardiac transcription factors [241], [242]. This is a particularly fascinating finding as it provides evidence that TBX5 is not only important for driving the expression of cardiac rhythm genes, but it is also driving the expression of transcriptional repressors that may prevent cardiac disease processes. While initial studies have defined target genes regulated by *Hopx*, they also suggest the possibility of independent functions. We also believe *Hopx* is not the only transcriptional repressor activated by TBX5, and understanding the different roles of these repressors will illuminate the gene regulatory network important for both cardiac homeostasis and pathology.

While our initial analyses were focused on the global trends and candidate genes identified in the time course, we also focused our interests on the differentially expressed signaling ligands at each timepoint. We believe examining the differentially expressed ligands may provide us with information regarding the important signaling pathways during the transition to atrial fibrillation. In Figure 4.S3, we have identified the top up-regulated and down-regulated ligands at each timepoint. These candidate ligands were derived from a left atria bulk RNA-seq, therefore we are not able to provide cell-specificity for the origin of these candidate ligands from this analysis. Fortunately, our laboratory has performed a single nuclear RNA-seq experiment in the left atria of control and *Tbx5*-deletion mice. This analysis is in the very preliminary stages, but it is clear that we will be able to provide clarity on the ligands and receptor interactions in the near future. For example, this will allow us to identify the cell-type from which the our candidate ligands are

expressed. We are particularly interested in *Bmp3* (bone morphogenetic protein 3), which is down-regulated from the earliest timepoint (D3), and is the most down-regulated ligands from Day 6 to Day 17. BMP3 is a member of the TGF-beta superfamily, and its role in the heart has not been examined. We find this to be a particularly fascinating candidate because a literature search in other tissues suggests it represses fibrosis and an inflammatory response. For example, decreased expression of BMP3 was shown in fibrotic lungs, and the presence of BMP3 reduced the activation of fibroblasts [245], [246] . Due to the fact that Bmp3 is significantly down-regulated at our earliest time point, we believe it is expressed in the cardiomyocytes during homeostasis. We hypothesize that the expression of BMP3 is necessary for homeostasis, and the down-regulation of BMP3 in cardiac disease leads to the cardiac pathology seen in atrial fibrillation. Again, at this point these are interesting candidates we have identified from our analysis and future studies are needed to validate or invalidate our hypotheses.

This timecourse RNA-seq was instrumental in understanding the time line of transcriptional changes that occur in atrial fibrillation. In the future, we hope to understand the function of these genes in homeostasis and how that is altered in cardiac disease models. We hope that understanding the effects of modulating these candidates will prevent insight into the pathophysiology that occurs due to atrial fibrillation, and potentially provide new avenues of therapeutic treatments.

Chapter 4.5 Acknowledgments

The bioinformatic analysis was performed in conjunction with Zhezhen Wang and Carlos Perez-Cervantes (Moskowitz Lab, University of Chicago).

CHAPTER 5: DISCUSSION

As the incidence of atrial fibrillation continues to increase worldwide, the elucidation of the molecular mechanisms driving the pathophysiology of disease are instrumental in identifying novel therapeutic treatments. As described in previous chapters, we utilized the *Tbx5*-deletion mouse model to provide instrumental insight into the various regulatory mechanisms of cardiac rhythm and atrial fibrillation. In Chapter 2, transcriptional profiling of small RNAs identified differentially expressed miRNAs in *Tbx5*-deficient mice. We utilized a high-throughput screen to identify miRNAs that have arrhythmogenic phenotypes, which we believe will be interesting candidates to pursue in future scientific endeavors. We also focused our subsequent efforts on a single candidate, miR-10b, and have shown that miR-10b expression is activated by TBX5 and is also a feedback mechanism to regulate TBX5 expression levels. In Chapter 3, we compared the transcriptional changes in the atrial fibrillation and heart failure mouse models to explore the shared pathophysiology between the two cardiac perturbations. They have a strong epidemiologic link, and their association has been attributed to their co-existence and shared risk factors. We utilized the transcriptional profiling of the coding and non-coding regions of the genome to identify transcriptional drivers and regulatory elements to understand their shared pathophysiology. Finally, in Chapter 4, a time course deletion of *Tbx5* was generated in an effort to identify the earliest transcriptional changes during the progression of the atrial fibrillation phenotype. This analysis was fruitful in identifying early transcriptional changes that are unique from the later transcriptional changes. The future scientific endeavors of our laboratory will be aimed at exploring the candidates from this study. The chapters in this body of work utilize the same *Tbx5*-deletion mouse model, but examine different aspects of the regulatory mechanisms in atrial fibrillation. Overall, we provide insight into the mechanisms regulating TBX5 expression

important for cardiac rhythm homeostasis, and the molecular changes that occur during cardiac disease.

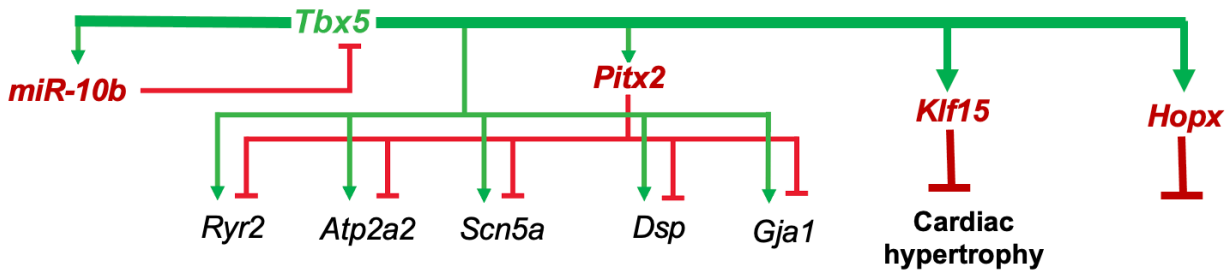


Figure 5.1. Expanded understanding of the *Tbx5* gene regulatory network. TBX5 and PITX2 drive expression of critical cardiac rhythm genes. TBX5 drives expression of miR-10b, which is a feedback mechanism to regulate the expression TBX5. TBX5 also drives expression of repressors, including *Klf15* and *Hopx*.

Cardiac rhythm is a dose sensitive process that requires precise gene expression to ensure proper cardiac function. Our small RNA transcriptional profiling study in the *Tbx5*-deletion mouse model was aimed at identifying changes in the miRNA regulatory landscape. We believe identifying differentially expressed miRNAs may identify candidates critical in maintaining cardiac rhythm homeostasis. The high-throughput screen allowed us to narrow down the list of candidates to those that affect the electrophysiologic signaling within the cardiomyocyte. It was fascinating that several candidates with arrhythmogenic phenotypes in the screen were previously validated to regulate the expression of cardiac rhythm genes, including miR-1231 and miR-21. We believe the remaining candidates warrant further studies to evaluate their function in the cardiomyocytes. In particular, we believe miR-1193 is a miRNA that plays an important role during a cardiac rhythm perturbation, and mis-expression of the miRNA affects the electrophysiologic signal in the cardiomyocyte. Identifying the target genes of miR-1193 will be the first critical step in identifying its putative functional role. Computational methods are a

important preliminary step in pairing the miRNAs with potential target genes. These bioinformatic tools are an invaluable resource, but their predictions require further validation. Transfection of miRNA mimics and inhibitors are a useful method to validate changes in target gene expression. After the transfection, quantitative PCR (qPCR) identifies changes in the RNA transcript levels and western blots identify changes in the protein levels. If the miRNA candidate is in fact regulating the expression of the candidate target gene, it can be validated using these methods. Additionally, luciferase reporter assays can be utilized to determine the repressive potential of the miRNA on a portion of the candidate gene, generally the 3' UTR. Mutations to the miRNA binding sites in the 3' UTR can provide further validation to an interaction between the miRNA and candidate target gene. Using these methods we would be able to provide insight into the genes targeted by this miRNA and their function in the heart. We hypothesize that miR-1193 is targeting genes important in cardiac rhythm, and alterations to miRNA expression will have an arrhythmogenic phenotype in the heart. We also believe elucidation into the miR-154 family will also provide meaningful insight into the pathophysiology of atrial fibrillation. Following the same steps described above will be important in determining the function of the miRNAs in the adult heart. These miRNAs are clustered together near and within *Mirg*. When considering potential gene targets, it would be insightful to consider predicted gene targets that are involved in the same pathway or biological process. Also identifying the transcriptional activators that drive the expression of miRNAs in these cardiac perturbations will be crucial to understanding the regulatory mechanisms of the miR-154 family. Future investigations will interrogate these exciting miRNA candidate that regulate cardiac rhythm gene expression.

In the adult heart, TBX5 is an important transcriptional activator of calcium-handling genes, demonstrating the importance of this transcription factor for stable cardiac rhythm.

Differential gene expression of the atrial fibrillation mouse model also revealed the activation of genes indicative of disease processes. Many of these remodeling processes in atrial fibrillation are shared with heart failure. Atrial fibrillation is considered an atrial disease, while heart failure is considered a ventricular disease, and yet there is a coalescence of the cardiac remodeling disease processes. Atrial fibrillation and heart failure have a strong epidemiologic link due to their shared incidence, shared risk factors, and shared cardiac remodeling processes. Therefore, our analysis was aimed at elucidating the shared transcriptional changes in these two cardiac perturbations in an effort to get a better understanding of the shared molecular mechanisms driving the pathophysiology of these cardiac diseases. Fascinatingly, the comparison of the coding and non-coding regions of the genome revealed a positive correlation of differential expression between the atrial fibrillation and heart failure mouse models. This demonstrated a shared transcriptional and regulatory landscape in both of these disease models. Our efforts were focused on the loci with shared differential ncRNA transcripts associated with shared differentially expressed transcription factors. Therefore, future endeavors in our laboratory will be aimed at elucidating the *Klf15* and *Sox9* loci. The binding of TBX5 at the promoter and enhancer of *Klf15* identified a novel role of TBX5 in driving the expression of a hypertrophy repressor. Future directions for this project include determining if the differential ncRNA transcript has a functional role in the regulation of *Klf15*, or whether the enhancer is the only functional requirement for *Klf15* expression. The transfection of antisense nucleotides (ASOs) targeting the ncRNA and polyA insertions to prematurely terminate transcription will be instrumental experiments to determine regulatory mechanisms of this ncRNA. We are also interested in exploring the TBX5-dependency of the human enhancer identified in the Human CAGE-seq dataset. Cloning the human enhancer into a luciferase reporter plasmid will allow us to validate its enhancer activity in cardiomyocytes, and

also examine the requirement of TBX5 with mutagenesis of the binding sites. Our ncRNA-seq analysis in atrial fibrillation and heart failure mouse model also identified shared up-regulated transcripts that are only expressed in the disease conditions. We focused our efforts on a ncRNA transcript near *Sox9*, a regulator of cardiac fibrosis. We have demonstrated that this transcript can be induced in an *in vitro* fibroblast experiment with the treatment of TGF-β. We hypothesize that this ncRNA transcript is activated during cardiac disease, and it plays a functional role in the fibroblasts during fibrotic processes. We speculate that repression of this ncRNA transcript will attenuate fibrosis during cardiac perturbations, including atrial fibrillation and heart failure. This a phenomenon that has been seen with the depletion of other ncRNAs in the mouse, including *Wisper* and MALAT1 [167], [168], [247]. We are also interested in exploring whether this ncRNA transcript is unique to cardiac fibroblasts or all fibrotic processes. We hypothesize that the exploration of other fibrotic tissues in publicly available datasets will demonstrate this presence of this ncRNA in those datasets. The comparison of atrial fibrillation and heart failure mouse models provided novel insights into the shared gene regulatory networks. Due to the fascinating findings and candidate loci identified in these comparisons, we have begun preliminary analysis on the overlap of *Tbx5*-deletion, TAC, and AngII (Angiotensin II) treated mice. Our collaborator, David Park, has previously demonstrated the similarities in the transcriptional response of TAC and AngII-treated mice [135]. Therefore, the integration of the coding and non-coding regions of the genome for all three datasets will identify transcriptional changes that are shared and unique to each cardiac perturbation, which may provide further understanding of cardiac disease processes.

Finally, the time course deletion of *Tbx5* allows our laboratory to elucidate the progression of the atrial fibrillation phenotype. The outcome of the RNA-seq time course shows that there is an earlier response confined to the changes in the cardiomyocytes, which progresses into

transcriptional changes into other cell types, including fibroblasts and immune cells. This distinction will allow us to identify changes in the cardiomyocytes that lead to potential paracrine signaling to the other cell types in the heart. This dataset has also provided us with an exciting group of candidates that may be responsible for the progression of atrial fibrillation pathophysiology. We were particularly focused on elucidating the earliest transcriptional changes, as those may be the inciting factor in the cardiac disease. For example, we hypothesize *Hopx* expression is important during homeostasis in the cardiomyocytes. In the atrial fibrillation and heart failure mouse models there is a down-regulation of the *Hopx* gene expression. This phenomenon is also seen in human heart failure, which was shown in the human heart failure CAGE-seq dataset. The change in expression begins at the Day 6 time point and continues to be differentially expressed throughout the time course. We believe *Hopx* is playing an important role in the maintaining cardiomyocyte homeostasis, and during cardiac stress *Hopx* expression levels are decreased early in the disease process. We believe *Hopx* is activated by TBX5 in a healthy heart based on the TBX5 ChIP-seq peaks identified at the promoter and nearby enhancer of *Hopx*. This suggests that TBX5 is driving the expression of a repressor that is inhibiting the activation of specific genes during cardiac homeostasis. Decreased expression levels of TBX5 will alleviate the repressive activity of *Hopx*. Closer examination of the developmental functions of *Hopx*, may provide meaningful insight into its role in the adult heart. For example, during development *Hopx* was shown to inhibit WNT signaling to promote myogenesis. We hypothesize that HOPX may be inhibiting the expression of Wnt ligands in the adult heart, and when *Hopx* expression is down-regulated during cardiac disease it alleviates this repression and leads to the activation of WNT signaling. WNT signaling is a well annotated signaling pathway in cardiac fibrosis, and this could potentially be a link connecting transcriptional changes in the cardiomyocytes that lead to

activation of cardiac fibroblasts [146]. While this is just speculation, future experiments will need to be completed to test our hypothesis. We believe further elucidating the functional role of *Hopx* in the heart will be instrumental in further understanding the transcriptional pathways that repress hypertrophy and cardiac fibrosis during homeostasis. Future experiments that can be used to begin examining this hypothesis would be to identify the binding sites of *Hopx* in the adult context, and identify molecular changes that occur in the adult heart with a conditional deletion mouse model. The generation of ChIP-seq genomic dataset for *Hopx* binding in the adult heart will allow us to determine if the genes bound by *Hopx* are different than those seen in a developmental context. Generating a conditional deletion of Hopx will allow us explore the transcriptional changes that occur in the adult heart in the absence of *Hopx*. While our hypothesis is pure speculation, it demonstrates the types of hypotheses we can generate and design future experiments around. We have now shown two examples of repressors, *Klf15* and *Hopx*, we believe to be activated by TBX5 in the wild-type state. However, they are also not the only TBX5-bound transcriptional repressor we have identified as differentially expressed in our RNA-seq time course dataset. TBX5 is canonically a transcriptional activator of downstream gene expression, but we believe that it is also driving the expression of a cohort of repressors. Elucidating the functional role of these repressors will provide insight into the indirect regulatory mechanisms of TBX5.

Overall, our work has leveraged the atrial fibrillation mouse model to elucidate the molecular mechanisms important for cardiac rhythm and cardiac disease. We utilized miRNAs, coding RNAs, and non-coding RNAs to build out the gene regulatory network driven by TBX5. Understanding the molecular mechanisms in driving the pathophysiology of the *Tbx5*-deletion mouse model will provide insight into the progression of atrial fibrillation, and potentially identify novel therapeutic approaches to manage the condition.

REFERENCES

[1] A. O. Grant, “Cardiac Ion Channels,” *Circulation Arrhythmia Electrophysiol*, vol. 2, no. 2, pp. 185–194, 2009, doi: 10.1161/circep.108.789081.

[2] B. T. Priest and J. S. McDermott, “Cardiac ion channels,” *Channels*, vol. 9, no. 6, pp. 352–359, 2015, doi: 10.1080/19336950.2015.1076597.

[3] H. Sutanto, A. Lyon, J. Lumens, U. Schotten, D. Dobrev, and J. Heijman, “Cardiomyocyte calcium handling in health and disease: Insights from in vitro and in silico studies,” *Prog Biophysics Mol Biology*, vol. 157, pp. 54–75, 2020, doi: 10.1016/j.pbiomolbio.2020.02.008.

[4] D. A. Eisner, J. L. Caldwell, K. Kistamás, and A. W. Trafford, “Calcium and Excitation-Contraction Coupling in the Heart,” *Circ Res*, vol. 121, no. 2, pp. 181–195, 2017, doi: 10.1161/circresaha.117.310230.

[5] C. J. Fearnley, H. L. Roderick, and M. D. Bootman, “Calcium signaling in cardiac myocytes.,” *Csh Perspect Biol*, vol. 3, no. 11, p. a004242, 2011, doi: 10.1101/cshperspect.a004242.

[6] S. L. Murphy, K. D. Kochanek, J. Xu, and E. Arias, “Mortality in the United States, 2020.,” *Nchs Data Brief*, no. 427, pp. 1–8, 2021.

[7] C. Antzelevitch and A. Burashnikov, “Overview of Basic Mechanisms of Cardiac Arrhythmia,” *Cardiac Electrophysiol Clin*, vol. 3, no. 1, pp. 23–45, 2011, doi: 10.1016/j.ccep.2010.10.012.

[8] A. H. Katsanos, H. Kamel, J. S. Healey, and R. G. Hart, “Stroke Prevention in Atrial Fibrillation,” *Circulation*, vol. 142, no. 24, pp. 2371–2388, 2020, doi: 10.1161/circulationaha.120.049768.

[9] G. F. Michaud and W. G. Stevenson, “Atrial Fibrillation,” *New Engl J Med*, vol. 384, no. 4, pp. 353–361, 2021, doi: 10.1056/nejmcp2023658.

[10] L. Staerk, J. A. Sherer, D. Ko, E. J. Benjamin, and R. H. Helm, “Atrial Fibrillation: Epidemiology, Pathophysiology, and Clinical Outcomes.,” *Circ Res*, vol. 120, no. 9, pp. 1501–1517, 2017, doi: 10.1161/circresaha.117.309732.

[11] A. S. Go *et al.*, “Prevalence of Diagnosed Atrial Fibrillation in Adults: National Implications for Rhythm Management and Stroke Prevention: the AnTicoagulation and Risk Factors In Atrial Fibrillation (ATRIA) Study,” *Jama*, vol. 285, no. 18, pp. 2370–2375, 2001, doi: 10.1001/jama.285.18.2370.

[12] Y. Miyasaka *et al.*, “Secular Trends in Incidence of Atrial Fibrillation in Olmsted County, Minnesota, 1980 to 2000, and Implications on the Projections for Future Prevalence,” *Circulation*, vol. 114, no. 2, pp. 119–125, 2006, doi: 10.1161/circulationaha.105.595140.

[13] S. S. Chugh *et al.*, “Worldwide Epidemiology of Atrial Fibrillation,” *Circulation*, vol. 129, no. 8, pp. 837–847, 2014, doi: 10.1161/circulationaha.113.005119.

[14] C. E. Woods and J. Olgin, “Atrial Fibrillation Therapy Now and in the Future,” *Circ Res*, vol. 114, no. 9, pp. 1532–1546, 2014, doi: 10.1161/circresaha.114.302362.

[15] J. Xu, J. G. Y. Luc, and K. Phan, “Atrial fibrillation: review of current treatment strategies,” *J Thorac Dis*, vol. 8, no. 9, pp. E886–E900, 2016, doi: 10.21037/jtd.2016.09.13.

[16] “Atrial fibrillation,” *Nat Rev Dis Primers*, vol. 8, no. 1, p. 20, 2022, doi: 10.1038/s41572-022-00354-w.

[17] L. M. Haegeli and H. Calkins, “Catheter ablation of atrial fibrillation: an update,” *Eur Heart J*, vol. 35, no. 36, pp. 2454–2459, 2014, doi: 10.1093/eurheartj/ehu291.

[18] Y. Iwasaki, K. Nishida, T. Kato, and S. Nattel, “Atrial Fibrillation Pathophysiology,” *Circulation*, vol. 124, no. 20, pp. 2264–2274, 2011, doi: 10.1161/circulationaha.111.019893.

[19] J. Heijman, N. Voigt, S. Nattel, and D. Dobrev, “Cellular and Molecular Electrophysiology of Atrial Fibrillation Initiation, Maintenance, and Progression,” *Circ Res*, vol. 114, no. 9, pp. 1483–1499, 2014, doi: 10.1161/circresaha.114.302226.

[20] C. Roselli, M. Rienstra, and P. T. Ellinor, “Genetics of Atrial Fibrillation in 2020: GWAS, Genome Sequencing, Polygenic Risk, and Beyond,” *Circ Res*, vol. 127, no. 1, pp. 21–33, 2020, doi: 10.1161/circresaha.120.316575.

[21] I. E. Christophersen and P. T. Ellinor, “Genetics of atrial fibrillation: from families to genomes,” *J Hum Genet*, vol. 61, no. 1, pp. 61–70, 2016, doi: 10.1038/jhg.2015.44.

[22] A. F. van Ouwerkerk *et al.*, “Epigenetic and Transcriptional Networks Underlying Atrial Fibrillation.,” *Circ. Res.*, vol. 127, no. 1, pp. 34–50, Jun. 2020, doi: 10.1161/circresaha.120.316574.

[23] N. R. Tucker and P. T. Ellinor, “Emerging Directions in the Genetics of Atrial Fibrillation,” *Circ Res*, vol. 114, no. 9, pp. 1469–1482, 2014, doi: 10.1161/circresaha.114.302225.

[24] S. A. Lubitz, C. Ozcan, J. W. Magnani, S. Kääb, E. J. Benjamin, and P. T. Ellinor, “Genetics of Atrial Fibrillation,” *Circulation Arrhythmia Electrophysiol*, vol. 3, no. 3, pp. 291–299, 2010, doi: 10.1161/circep.110.942441.

[25] J. Feghaly, P. Zakka, B. London, C. A. MacRae, and M. M. Refaat, “Genetics of Atrial Fibrillation,” *J Am Hear Assoc Cardiovasc Cerebrovasc Dis*, vol. 7, no. 20, p. e009884, 2018, doi: 10.1161/jaha.118.009884.

[26] P. T. Ellinor *et al.*, “Common variants in KCNN3 are associated with lone atrial fibrillation,” *Nat Genet*, vol. 42, no. 3, pp. 240–244, 2010, doi: 10.1038/ng.537.

[27] A. Pfeufer *et al.*, “Genome-wide association study of PR interval,” *Nat Genet*, vol. 42, no. 2, pp. 153–159, 2010, doi: 10.1038/ng.517.

[28] P. T. Ellinor *et al.*, “Meta-analysis identifies six new susceptibility loci for atrial fibrillation,” *Nat Genet*, vol. 44, no. 6, pp. 670–675, 2012, doi: 10.1038/ng.2261.

[29] X. Zang *et al.*, “SNP rs3825214 in TBX5 Is Associated with Lone Atrial Fibrillation in Chinese Han Population,” *Plos One*, vol. 8, no. 5, p. e64966, 2013, doi: 10.1371/journal.pone.0064966.

[30] D. F. Gudbjartsson *et al.*, “Variants conferring risk of atrial fibrillation on chromosome 4q25,” *Nature*, vol. 448, no. 7151, pp. 353–357, 2007, doi: 10.1038/nature06007.

[31] E. J. Benjamin *et al.*, “Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry,” *Nat Genet*, vol. 41, no. 8, pp. 879–881, 2009, doi: 10.1038/ng.416.

[32] S. Chen *et al.*, “KCNQ1 mutations in patients with a family history of lethal cardiac arrhythmias and sudden death,” *Clin Genet*, vol. 63, no. 4, pp. 273–282, 2003, doi: 10.1034/j.1399-0004.2003.00048.x.

[33] D. C. BARTOS *et al.*, “A KCNQ1 Mutation Causes a High Penetrance for Familial Atrial Fibrillation,” *J Cardiovasc Electr*, vol. 24, no. 5, pp. 562–569, 2013, doi: 10.1111/jce.12068.

[34] C.-S. Ki *et al.*, “A KCNQ1 mutation causes age-dependant bradycardia and persistent atrial fibrillation,” *Pflügers Archiv - European J Physiology*, vol. 466, no. 3, pp. 529–540, 2014, doi: 10.1007/s00424-013-1337-6.

[35] K. Hasegawa *et al.*, “A novel KCNQ1 missense mutation identified in a patient with juvenile-onset atrial fibrillation causes constitutively open IKs channels,” *Heart Rhythm*, vol. 11, no. 1, pp. 67–75, 2014, doi: 10.1016/j.hrthm.2013.09.073.

[36] J. Wu *et al.*, “A trafficking-deficient KCNQ1 mutation, T587M, causes a severe phenotype of long QT syndrome by interfering with intracellular hERG transport,” *J Cardiol*, vol. 73, no. 5, pp. 343–350, 2019, doi: 10.1016/j.jjcc.2018.10.011.

[37] A. Lundby, L. S. Ravn, J. H. Svendsen, S.-P. Olesen, and N. Schmitt, “KCNQ1 mutation Q147R is associated with atrial fibrillation and prolonged QT interval,” *Heart Rhythm*, vol. 4, no. 12, pp. 1532–1541, 2007, doi: 10.1016/j.hrthm.2007.07.022.

[38] S. Das *et al.*, “Mutation in the S3 segment of KCNQ1 results in familial lone atrial fibrillation,” *Heart Rhythm*, vol. 6, no. 8, pp. 1146–1153, 2009, doi: 10.1016/j.hrthm.2009.04.015.

[39] D. C. Bartos *et al.*, “R231C mutation in KCNQ1 causes long QT syndrome type 1 and familial atrial fibrillation.,” *Heart Rhythm*, vol. 8, no. 1, pp. 48–55, 2010, doi: 10.1016/j.hrthm.2010.09.010.

[40] M. S. Olesen *et al.*, “A novel KCND3 gain-of-function mutation associated with early-onset of persistent lone atrial fibrillation,” *Cardiovasc Res*, vol. 98, no. 3, pp. 488–495, 2013, doi: 10.1093/cvr/cvt028.

[41] T. Makiyama *et al.*, “A Novel SCN5A Gain-of-Function Mutation M1875T Associated With Familial Atrial Fibrillation,” *J Am Coll Cardiol*, vol. 52, no. 16, pp. 1326–1334, 2008, doi: 10.1016/j.jacc.2008.07.013.

[42] I. E. Christophersen *et al.*, “Genetic variation in KCNA5: impact on the atrial-specific potassium current IKur in patients with lone atrial fibrillation,” *Eur Heart J*, vol. 34, no. 20, pp. 1517–1525, 2013, doi: 10.1093/eurheartj/ehs442.

[43] D. Darbar *et al.*, “Cardiac Sodium Channel (SCN5A) Variants Associated with Atrial Fibrillation,” *Circulation*, vol. 117, no. 15, pp. 1927–1935, 2008, doi: 10.1161/circulationaha.107.757955.

[44] P. T. Ellinor, E. G. Nam, M. A. Shea, D. J. Milan, J. N. Ruskin, and C. A. MacRae, “Cardiac sodium channel mutation in atrial fibrillation,” *Heart Rhythm*, vol. 5, no. 1, pp. 99–105, 2008, doi: 10.1016/j.hrthm.2007.09.015.

[45] T. M. Olson *et al.*, “Kv1.5 channelopathy due to KCNA5 loss-of-function mutation causes human atrial fibrillation,” *Hum Mol Genet*, vol. 15, no. 14, pp. 2185–2191, 2006, doi: 10.1093/hmg/ddl143.

[46] M. Deo *et al.*, “KCNJ2 mutation in short QT syndrome 3 results in atrial fibrillation and ventricular proarrhythmia,” *Proc National Acad Sci*, vol. 110, no. 11, pp. 4291–4296, 2013, doi: 10.1073/pnas.1218154110.

[47] Y. Tao *et al.*, “Pitx2, an Atrial Fibrillation Predisposition Gene, Directly Regulates Ion Transport and Intercalated Disc Genes,” *Circulation Cardiovasc Genetics*, vol. 7, no. 1, pp. 23–32, 2018, doi: 10.1161/circgenetics.113.000259.

[48] P. Kirchhof *et al.*, “PITX2c Is Expressed in the Adult Left Atrium, and Reducing Pitx2c Expression Promotes Atrial Fibrillation Inducibility and Complex Changes in Gene Expression,” *Circulation Cardiovasc Genetics*, vol. 4, no. 2, pp. 123–133, 2011, doi: 10.1161/circgenetics.110.958058.

[49] A. Chinchilla *et al.*, “PITX2 Insufficiency Leads to Atrial Electrical and Structural Remodeling Linked to Arrhythmogenesis,” *Circulation Cardiovasc Genetics*, vol. 4, no. 3, pp. 269–279, 2011, doi: 10.1161/circgenetics.110.958116.

[50] C. Roselli *et al.*, “Multi-ethnic genome-wide association study for atrial fibrillation.,” *Nature Genetics*, vol. 50, no. 9, pp. 1225–1233, 2018, doi: 10.1038/s41588-018-0133-9.

[51] J. B. Nielsen *et al.*, “Genome-wide Study of Atrial Fibrillation Identifies Seven Risk Loci and Highlights Biological Pathways and Regulatory Elements Involved in Cardiac Development,” *Am J Hum Genetics*, vol. 102, no. 1, pp. 103–115, 2018, doi: 10.1016/j.ajhg.2017.12.003.

[52] R. B. Thorolfsdottir *et al.*, “Coding variants in RPL3L and MYZAP increase risk of atrial fibrillation,” *Commun Biology*, vol. 1, no. 1, p. 68, 2018, doi: 10.1038/s42003-018-0068-9.

[53] I. E. Christophersen *et al.*, “Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation,” *Nat Genet*, vol. 49, no. 6, pp. 946–952, 2017, doi: 10.1038/ng.3843.

[54] Afg. Consortium *et al.*, “Identification of six new genetic loci associated with atrial fibrillation in the Japanese population,” *Nat Genet*, vol. 49, no. 6, pp. 953–958, 2017, doi: 10.1038/ng.3842.

[55] M. F. Sinner *et al.*, “Integrating Genetic, Transcriptional, and Functional Analyses to Identify 5 Novel Genes for Atrial Fibrillation,” *Circulation*, vol. 130, no. 15, pp. 1225–1235, 2014, doi: 10.1161/circulationaha.114.009892.

[56] D. F. Gudbjartsson *et al.*, “A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke,” *Nat Genet*, vol. 41, no. 8, pp. 876–878, 2009, doi: 10.1038/ng.417.

[57] J.-Y. Lee *et al.*, “Korean atrial fibrillation network genome-wide association study for early-onset atrial fibrillation identifies novel susceptibility loci,” *Eur Heart J*, vol. 38, no. 34, pp. 2586–2594, 2017, doi: 10.1093/eurheartj/ehx213.

[58] B. N. Akerberg *et al.*, “A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers,” *Nat Commun*, vol. 10, no. 1, p. 4907, 2019, doi: 10.1038/s41467-019-12812-3.

[59] M. Nadeau *et al.*, “An endocardial pathway involving Tbx5, Gata4, and Nos3 required for atrial septum formation,” *Proc National Acad Sci*, vol. 107, no. 45, pp. 19356–19361, 2010, doi: 10.1073/pnas.0914888107.

[60] L. Luna-Zurita *et al.*, “Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis.,” *Cell*, vol. 164, no. 5, pp. 999–1014, Feb. 2016, doi: 10.1016/j.cell.2016.01.004.

[61] M. Maitra *et al.*, “Interaction of Gata4 and Gata6 with Tbx5 is critical for normal cardiac development,” *Dev Biol*, vol. 326, no. 2, pp. 368–377, 2009, doi: 10.1016/j.ydbio.2008.11.004.

[62] V. L. F. Linhares *et al.*, “Transcriptional regulation of the murine Connexin40 promoter by cardiac factors Nkx2-5, GATA4 and Tbx5,” *Cardiovasc Res*, vol. 64, no. 3, pp. 402–411, 2004, doi: 10.1016/j.cardiores.2004.09.021.

[63] I. P. G. Moskowitz *et al.*, “A molecular pathway including Id2, Tbx5, and Nkx2-5 required for cardiac conduction system development.,” *Cell*, vol. 129, no. 7, pp. 1365–1376, Jun. 2007, doi: 10.1016/j.cell.2007.04.036.

[64] B. G. Bruneau *et al.*, “A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease.,” *Cell*, vol. 106, no. 6, pp. 709–721, Sep. 2001, doi: 10.1016/s0092-8674(01)00493-7.

[65] V. Garg *et al.*, “GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5,” *Nature*, vol. 424, no. 6947, pp. 443–447, 2003, doi: 10.1038/nature01827.

[66] Y.-S. Ang *et al.*, “Disease Model of GATA4 Mutation Reveals Transcription Factor Cooperativity in Human Cardiogenesis,” *Cell*, vol. 167, no. 7, pp. 1734-1749.e22, 2016, doi: 10.1016/j.cell.2016.11.033.

[67] C. Misra, S.-W. Chang, M. Basu, N. Huang, and V. Garg, “Disruption of myocardial Gata4 and Tbx5 results in defects in cardiomyocyte proliferation and atrioventricular septation,” *Hum Mol Genet*, vol. 23, no. 19, pp. 5025–5035, 2014, doi: 10.1093/hmg/ddu215.

[68] L. Pradhan *et al.*, “Intermolecular Interactions of Cardiac Transcription Factors NKX2.5 and TBX5,” *Biochemistry-us*, vol. 55, no. 12, pp. 1702–1710, 2016, doi: 10.1021/acs.biochem.6b00171.

[69] A. He, S. W. Kong, Q. Ma, and W. T. Pu, “Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart,” *Proc National Acad Sci*, vol. 108, no. 14, pp. 5632–5637, 2011, doi: 10.1073/pnas.1016959108.

[70] Y.-Q. Yang *et al.*, “GATA4 loss-of-function mutations in familial atrial fibrillation,” *Clin Chim Acta*, vol. 412, no. 19–20, pp. 1825–1830, 2011, doi: 10.1016/j.cca.2011.06.017.

[71] J.-Q. Jiang, F.-F. Shen, W.-Y. Fang, X. Liu, and Y.-Q. Yang, “Novel GATA4 mutations in lone atrial fibrillation,” *Int J Mol Med*, vol. 28, no. 6, pp. 1025–32, 2011, doi: 10.3892/ijmm.2011.783.

[72] R.-T. HUANG, S. XUE, Y.-J. XU, M. ZHOU, and Y.-Q. YANG, “A novel NKX2.5 loss-of-function mutation responsible for familial atrial fibrillation,” *Int J Mol Med*, vol. 31, no. 5, pp. 1119–1126, 2013, doi: 10.3892/ijmm.2013.1316.

[73] W.-H. Xie *et al.*, “Prevalence and spectrum of Nkx2.5 mutations associated with idiopathic atrial fibrillation,” *Clinics*, vol. 68, no. 6, pp. 777–784, 2013, doi: 10.6061/clinics/2013(06)09.

[74] M. G. Posch *et al.*, “Mutations in the cardiac transcription factor GATA4 in patients with lone atrial fibrillation,” *Eur J Med Genet*, vol. 53, no. 4, pp. 201–203, 2010, doi: 10.1016/j.ejmg.2010.03.008.

[75] J. Wang, Y.-M. Sun, and Y.-Q. Yang, “Mutation spectrum of the GATA4 gene in patients with idiopathic atrial fibrillation,” *Mol Biol Rep*, vol. 39, no. 8, pp. 8127–8135, 2012, doi: 10.1007/s11033-012-1660-6.

[76] H. Holm *et al.*, “Several common variants modulate heart rate, PR interval and QRS duration,” *Nat Genet*, vol. 42, no. 2, pp. 117–122, 2010, doi: 10.1038/ng.511.

[77] J.-F. Ma *et al.*, “TBX5 mutations contribute to early-onset atrial fibrillation in Chinese and Caucasians,” *Cardiovasc Res*, vol. 109, no. 3, pp. 442–450, 2016, doi: 10.1093/cvr/cvw003.

[78] Z.-C. Wang *et al.*, “Prevalence and Spectrum of TBX5 Mutation in Patients with Lone Atrial Fibrillation,” *Int J Med Sci*, vol. 13, no. 1, pp. 60–67, 2016, doi: 10.7150/ijms.13264.

[79] M. Holt and S. Oram, “FAMILIAL HEART DISEASE WITH SKELETAL MALFORMATIONS,” *Brit Heart J*, vol. 22, no. 2, p. 236, 1960, doi: 10.1136/hrt.22.2.236.

[80] R. A. Newbury-Ecob, R. Leanage, J. A. Raeburn, and I. D. Young, “Holt-Oram syndrome: a clinical genetic study.,” *J Med Genet*, vol. 33, no. 4, p. 300, 1996, doi: 10.1136/jmg.33.4.300.

[81] C. T. Basson *et al.*, “The Clinical and Genetic Spectrum of the Holt-Oram Syndrome (Heart-Hand Syndrome),” *New Engl J Medicine*, vol. 330, no. 13, pp. 885–891, 1994, doi: 10.1056/nejm199403313301302.

[82] C. T. Basson *et al.*, “Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome.,” *Nat Genet*, vol. 15, no. 1, pp. 30–5, 1997, doi: 10.1038/ng0197-30.

[83] Q. Y. Li *et al.*, “Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family,” *Nat Genet*, vol. 15, no. 1, pp. 21–29, 1997, doi: 10.1038/ng0197-21.

[84] A. V. Postma *et al.*, “A Gain-of-Function TBX5 Mutation Is Associated With Atypical Holt–Oram Syndrome and Paroxysmal Atrial Fibrillation,” *Circ Res*, vol. 102, no. 11, pp. 1433–1442, 2008, doi: 10.1161/circresaha.107.168294.

[85] V. Wilson and F. L. Conlon, “The T-box family,” *Genome Biol*, vol. 3, no. 6, p. reviews3008.1, 2002, doi: 10.1186/gb-2002-3-6-reviews3008.

[86] V. E. Papaioannou, “The T-box gene family: emerging roles in development, stem cells and cancer,” *Development*, vol. 141, no. 20, pp. 3819–3833, 2014, doi: 10.1242/dev.104471.

[87] V. E. Papaioannou and L. M. Silver, “The T-box gene family,” *Bioessays*, vol. 20, no. 1, pp. 9–19, 1998, doi: 10.1002/(sici)1521-1878(199801)20:1<9::aid-bies4>3.0.co;2-q.

[88] B. G. Herrmann, S. Labeit, A. Poustka, T. R. King, and H. Lehrach, “Cloning of the T gene required in mesoderm formation in the mouse,” *Nature*, vol. 343, no. 6259, pp. 617–622, 1990, doi: 10.1038/343617a0.

[89] A. Kispert and B. G. Herrmann, “The Brachyury gene encodes a novel DNA binding protein.,” *Embo J*, vol. 12, no. 8, pp. 3211–3220, 1993, doi: 10.1002/j.1460-2075.1993.tb05990.x.

[90] A. Kispert, B. Koschorz, and B. G. Herrmann, “The T protein encoded by Brachyury is a tissue-specific transcription factor.,” *Embo J*, vol. 14, no. 19, pp. 4763–72, 1995, doi: 10.1002/j.1460-2075.1995.tb00158.x.

[91] L. A. Naiche, Z. Harrelson, R. G. Kelly, and V. E. Papaioannou, “T-BOX GENES IN VERTEBRATE DEVELOPMENT,” *Genetics*, vol. 39, no. 1, pp. 219–239, 2005, doi: 10.1146/annurev.genet.39.073003.105925.

[92] S. I. Agulnik *et al.*, “Evolution of Mouse T-box Genes by Tandem Duplication and Cluster Dispersion,” *Genetics*, vol. 144, no. 1, pp. 249–254, 1996, doi: 10.1093/genetics/144.1.249.

[93] I. Ruvinsky and L. M. Silver, “Newly Identified Paralogous Groups on Mouse Chromosomes 5 and 11 Reveal the Age of a T-Box Cluster Duplication,” *Genomics*, vol. 40, no. 2, pp. 262–266, 1997, doi: 10.1006/geno.1996.4591.

[94] A. Calmont *et al.*, “Tbx1 controls cardiac neural crest cell migration during arch artery development by regulating Gbx2 expression in the pharyngeal ectoderm,” *Development*, vol. 136, no. 18, pp. 3173–3183, 2009, doi: 10.1242/dev.028902.

[95] E. A. Lindsay *et al.*, “Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice,” *Nature*, vol. 410, no. 6824, pp. 97–101, 2001, doi: 10.1038/35065105.

[96] L. Chen, F. G. Fulcoli, S. Tang, and A. Baldini, “Tbx1 Regulates Proliferation and Differentiation of Multipotent Heart Progenitors,” *Circ Res*, vol. 105, no. 9, pp. 842–851, 2009, doi: 10.1161/circresaha.109.200295.

[97] L. A. Jerome and V. E. Papaioannou, “DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1,” *Nat Genet*, vol. 27, no. 3, pp. 286–291, 2001, doi: 10.1038/85845.

[98] P. J. Scambler, “The 22q11 deletion syndromes,” *Hum Mol Genet*, vol. 9, no. 16, pp. 2421–2426, 2000, doi: 10.1093/hmg/9.16.2421.

[99] F. Greulich, C. Rudat, and A. Kispert, “Mechanisms of T-box gene function in the developing heart,” *Cardiovasc Res*, vol. 91, no. 2, pp. 212–222, 2011, doi: 10.1093/cvr/cvr112.

[100] W. T. J. Aanhaanen *et al.*, “Defective Tbx2-dependent patterning of the atrioventricular canal myocardium causes accessory pathway formation in mice,” *J Clin Invest*, vol. 121, no. 2, pp. 534–544, 2011, doi: 10.1172/jci44350.

[101] Z. Harrelson *et al.*, “Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development,” *Development*, vol. 131, no. 20, pp. 5041–5052, 2004, doi: 10.1242/dev.01378.

[102] V. M. Christoffels, W. M. H. Hoogaars, A. Tessari, D. E. W. Clout, A. F. M. Moorman, and M. Campione, “T-box transcription factor Tbx2 represses differentiation and formation of the cardiac chambers,” *Dev Dynam*, vol. 229, no. 4, pp. 763–770, 2004, doi: 10.1002/dvdy.10487.

[103] M. L. Bakker *et al.*, “Transcription Factor Tbx3 Is Required for the Specification of the Atrioventricular Conduction System,” *Circ Res*, vol. 102, no. 11, pp. 1340–1349, 2008, doi: 10.1161/circresaha.107.169565.

[104] C. Liu, A. Shen, X. Li, W. Jiao, X. Zhang, and Z. Li, “T-box transcription factor TBX20 mutations in Chinese patients with congenital heart disease,” *Eur J Med Genet*, vol. 51, no. 6, pp. 580–587, 2008, doi: 10.1016/j.ejmg.2008.09.001.

[105] M. K. Singh *et al.*, “Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2,” *Development*, vol. 132, no. 12, pp. 2697–2707, 2005, doi: 10.1242/dev.01854.

[106] E. P. Kirk *et al.*, “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 Genetics*, vol. 81, no. 2, pp. 280–291, 2007, doi: 10.1086/519530.

[107] F. A. Stennard *et al.*, “Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation,” *Development*, vol. 132, no. 10, pp. 2451–2462, 2005, doi: 10.1242/dev.01799.

[108] B. G. Bruneau *et al.*, “Chamber-Specific Cardiac Expression of Tbx5 and Heart Defects in Holt–Oram Syndrome,” *Dev Biol*, vol. 211, no. 1, pp. 100–108, 1999, doi: 10.1006/dbio.1999.9298.

[109] D. L. Chapman *et al.*, “Expression of the T-box family genes, Tbx1–Tbx5, during early mouse development,” *Dev Dynam*, vol. 206, no. 4, pp. 379–390, 1996, doi: 10.1002/(sici)1097-0177(199608)206:4<379::aid-aja4>3.0.co;2-f.

[110] J. J. Gibson-Brown, S. I. Agulnik, L. M. Silver, and V. E. Papaioannou, “Expression of T-box genes Tbx2–Tbx5 during chick organogenesis,” *Mech Develop*, vol. 74, no. 1–2, pp. 165–169, 1998, doi: 10.1016/s0925-4773(98)00056-2.

[111] J. J. Gibson-Brown, S. I. Agulnik, L. M. Silver, L. Niswander, and V. E. Papaioannou, “Involvement of T-box genes Tbx2-Tbx5 in vertebrate limb specification and development.,” *Dev Camb Engl*, vol. 125, no. 13, pp. 2499–509, 1998, doi: 10.1242/dev.125.13.2499.

[112] J. J. Gibson-Brown *et al.*, “Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity.,” *Mech Develop*, vol. 56, no. 1–2, pp. 93–101, 1996, doi: 10.1016/0925-4773(96)00514-x.

[113] A. Isaac *et al.*, “Tbx genes and limb identity in chick embryo development,” *Development*, vol. 125, no. 10, pp. 1867–1875, 1998, doi: 10.1242/dev.125.10.1867.

[114] P. Agarwal *et al.*, “Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo,” *Development*, vol. 130, no. 3, pp. 623–633, 2003, doi: 10.1242/dev.00191.

[115] I. P. G. Moskowitz *et al.*, “The T-Box transcription factor Tbx5 is required for the patterning and maturation of the murine cardiac conduction system.,” *Development*, vol. 131, no. 16, pp. 4107–4116, 2004, doi: 10.1242/dev.01265.

[116] D. E. Arnolds *et al.*, “TBX5 drives Scn5a expression to regulate cardiac conduction system function.,” *J. Clin. Invest.*, vol. 122, no. 7, pp. 2509–2518, 2012, doi: 10.1172/jci62617.

[117] O. Burnicka-Turek *et al.*, “Transcriptional Patterning of the Ventricular Cardiac Conduction System,” *Circ Res*, vol. 127, no. 3, pp. e94–e106, 2020, doi: 10.1161/circresaha.118.314460.

[118] R. D. Nadadur *et al.*, “Pitx2 modulates a Tbx5-dependent gene regulatory network to maintain atrial rhythm.,” *Sci Transl Med*, vol. 8, no. 354, pp. 354ra115-354ra115, Aug. 2016, doi: 10.1126/scitranslmed.aaf4891.

[119] B. Laforest *et al.*, “Atrial fibrillation risk loci interact to modulate Ca2+-dependent atrial rhythm homeostasis.,” *J. Clin. Invest.*, vol. 129, no. 11, pp. 4937–4950, Nov. 2019, doi: 10.1172/jci124231.

[120] W. Dai *et al.*, “A calcium transport mechanism for atrial fibrillation in Tbx5-mutant mice.,” *Elife*, vol. 8, p. 40, Mar. 2019, doi: 10.7554/elife.41814.

[121] C. W. Tsao *et al.*, “Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association,” *Circulation*, vol. 145, no. 8, pp. e153–e639, doi: 10.1161/cir.0000000000001052.

[122] T. J. Wang *et al.*, “Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study.,” *Circulation*, vol. 107, no. 23, pp. 2920–5, 2003, doi: 10.1161/01.cir.0000072767.89944.6e.

[123] J. S. Burchfield, M. Xie, and J. A. Hill, “Pathological Ventricular Remodeling: Mechanisms: Part 1 of 2,” *Circulation*, vol. 128, no. 4, pp. 388–400, 2013, doi: 10.1161/circulationaha.113.001878.

[124] S. J. Simmonds, I. Cuijpers, S. Heymans, and E. A. V. Jones, “Cellular and Molecular Differences between HFpEF and HFrEF: A Step Ahead in an Improved Pathological Understanding,” *Cells*, vol. 9, no. 1, p. 242, 2020, doi: 10.3390/cells9010242.

[125] G. Savarese, D. of C. Sweden Department of Medicine, Karolinska Insitutet, Stockholm, D. of C. Sweden Karolinska University Hospital, Stockholm, and L. H. Lund, “Global Public Health Burden of Heart Failure,” *Cardiac Fail Rev*, vol. 03, no. 01, p. 7, 2017, doi: 10.15420/cfr.2016:25:2.

[126] A. L. Bui, T. B. Horwich, and G. C. Fonarow, “Epidemiology and risk profile of heart failure,” *Nat Rev Cardiol*, vol. 8, no. 1, pp. 30–41, 2011, doi: 10.1038/nrcardio.2010.165.

[127] P. A. Heidenreich *et al.*, “Forecasting the Impact of Heart Failure in the United States,” *Circulation Hear Fail*, vol. 6, no. 3, pp. 606–619, 2018, doi: 10.1161/hhf.0b013e318291329a.

[128] R. D. Patten and M. R. Hall-Porter, “Small Animal Models of Heart Failure,” *Circulation Hear Fail*, vol. 2, no. 2, pp. 138–144, 2009, doi: 10.1161/circheartfailure.108.839761.

[129] C. Riehle and J. Bauersachs, “Small animal models of heart failure,” *Cardiovasc Res*, vol. 115, no. 13, pp. 1838–1849, 2019, doi: 10.1093/cvr/cvz161.

[130] P. M. Pilz *et al.*, “Large and Small Animal Models of Heart Failure With Reduced Ejection Fraction,” *Circ Res*, vol. 130, no. 12, pp. 1888–1905, 2022, doi: 10.1161/circresaha.122.320246.

[131] R. Tavakoli, S. Nemska, P. Jamshidi, M. Gassmann, and N. Frossard, “Technique of Minimally Invasive Transverse Aortic Constriction in Mice for Induction of Left Ventricular Hypertrophy.,” *J Vis Exp Jove*, no. 127, 2017, doi: 10.3791/56231.

[132] H. A. Rockman, S. P. Wachhorst, L. Mao, and J. Ross, “ANG II receptor blockade prevents ventricular hypertrophy and ANF gene expression with pressure overload in mice,” *Am J Physiol-heart C*, vol. 266, no. 6, pp. H2468–H2475, 1994, doi: 10.1152/ajpheart.1994.266.6.h2468.

[133] D. A. Richards *et al.*, “Distinct Phenotypes Induced by Three Degrees of Transverse Aortic Constriction in Mice,” *Sci Rep-uk*, vol. 9, no. 1, p. 5844, 2019, doi: 10.1038/s41598-019-42209-7.

[134] M. Alexanian *et al.*, “A transcriptional switch governs fibroblast activation in heart disease,” *Nature*, vol. 595, no. 7867, pp. 438–443, 2021, doi: 10.1038/s41586-021-03674-1.

[135] N. Yamaguchi *et al.*, “Cardiac Pressure Overload Decreases ETV1 Expression in the Left Atrium, Contributing to Atrial Electrical and Structural Remodeling,” *Circulation*, vol. 143, no. 8, pp. 805–820, 2021, doi: 10.1161/circulationaha.120.048121.

[136] C. Riehle *et al.*, “PGC-1β Deficiency Accelerates the Transition to Heart Failure in Pressure Overload Hypertrophy,” *Circ Res*, vol. 109, no. 7, pp. 783–793, 2011, doi: 10.1161/circresaha.111.243964.

[137] A. Grund *et al.*, “A gene therapeutic approach to inhibit calcium and integrin binding protein 1 ameliorates maladaptive remodelling in pressure overload,” *Cardiovasc Res*, vol. 115, no. 1, pp. 71–82, 2018, doi: 10.1093/cvr/cvy154.

[138] N. Frey, H. A. Katus, E. N. Olson, and J. A. Hill, “Hypertrophy of the Heart,” *Circulation*, vol. 109, no. 13, pp. 1580–1589, 2004, doi: 10.1161/01.cir.0000120390.68287.bb.

[139] M. Samak *et al.*, “Cardiac Hypertrophy: An Introduction to Molecular and Cellular Basis,” *Medical Sci Monit Basic Res*, vol. 22, pp. 75–79, 2016, doi: 10.12659/msmbr.900437.

[140] J. Heineke and J. D. Molkentin, “Regulation of cardiac hypertrophy by intracellular signalling pathways,” *Nat Rev Mol Cell Bio*, vol. 7, no. 8, pp. 589–600, 2006, doi: 10.1038/nrm1983.

[141] H. Akazawa and I. Komuro, “Roles of Cardiac Transcription Factors in Cardiac Hypertrophy,” *Circulation Res J Am Hear Assoc*, vol. 92, no. 10, pp. 1079–1088, 2003, doi: 10.1161/01.res.0000072977.86706.23.

[142] T. P. Creamer, “Calcineurin,” *Cell Commun Signal*, vol. 18, no. 1, p. 137, 2020, doi: 10.1186/s12964-020-00636-4.

[143] G. Krenning, E. M. Zeisberg, and R. Kalluri, “The origin of fibroblasts and mechanism of cardiac fibrosis,” *J Cell Physiol*, vol. 225, no. 3, pp. 631–637, 2010, doi: 10.1002/jcp.22322.

[144] J. G. Travers, F. A. Kamal, J. Robbins, K. E. Yutzey, and B. C. Blaxall, “Cardiac Fibrosis: The Fibroblast Awakens.,” *Circ. Res.*, vol. 118, no. 6, pp. 1021–1040, Mar. 2016, doi: 10.1161/circresaha.115.306565.

[145] H. Khalil *et al.*, “Fibroblast-specific TGF-β–Smad2/3 signaling underlies cardiac fibrosis,” *J Clin Invest*, vol. 127, no. 10, pp. 3770–3783, 2017, doi: 10.1172/jci94753.

[146] F. Yousefi *et al.*, “TGF-β and WNT signaling pathways in cardiac fibrosis: non-coding RNAs come into focus,” *Cell Commun Signal Ccs*, vol. 18, no. 1, p. 87, 2020, doi: 10.1186/s12964-020-00555-4.

[147] S. Mahida *et al.*, “Overexpression of KCNN3 results in sudden cardiac death,” *Cardiovasc Res*, vol. 101, no. 2, pp. 326–334, 2013, doi: 10.1093/cvr/cvt269.

[148] J. Temple *et al.*, “Atrial Fibrillation in KCNE1-Null Mice,” *Circ Res*, vol. 97, no. 1, pp. 62–69, 2005, doi: 10.1161/01.res.0000173047.42236.88.

[149] J. O’Brien, H. Hayder, Y. Zayed, and C. Peng, “Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation,” *Front Endocrinol*, vol. 9, p. 402, 2018, doi: 10.3389/fendo.2018.00402.

[150] Y. Cai, X. Yu, S. Hu, and J. Yu, “A brief review on the mechanisms of miRNA regulation.,” *Genomics Proteomics Bioinformatics*, vol. 7, no. 4, pp. 147–154, 2009, doi: 10.1016/s1672-0229(08)60044-3.

[151] M. Ha and V. N. Kim, “Regulation of microRNA biogenesis.,” *Nature Publishing Group*, vol. 15, no. 8, pp. 509–524, 2014, doi: 10.1038/nrm3838.

[152] L. F. R. Gebert and I. J. MacRae, “Regulation of microRNA function in animals.,” *Nature Publishing Group*, vol. 21, p. 743, Aug. 2018, doi: 10.1038/s41580-018-0045-7.

[153] D. P. Bartel, “MicroRNAs Genomics, Biogenesis, Mechanism, and Function,” *Cell*, vol. 116, no. 2, pp. 281–297, 2004, doi: 10.1016/s0092-8674(04)00045-5.

[154] K. R. Cordes and D. Srivastava, “MicroRNA Regulation of Cardiovascular Development,” *Circ Res*, vol. 104, no. 6, pp. 724–732, 2009, doi: 10.1161/circresaha.108.192872.

[155] G. H. Kim, “MicroRNA regulation of cardiac conduction and arrhythmias.,” *Transl Res*, vol. 161, no. 5, pp. 381–392, 2013, doi: 10.1016/j.trsl.2012.12.004.

[156] K. Saliminejad, H. R. K. Khorshid, S. S. Fard, and S. H. Ghaffari, “An overview of microRNAs: Biology, functions, therapeutics, and analysis methods,” *J Cell Physiol*, vol. 234, no. 5, pp. 5451–5465, 2019, doi: 10.1002/jcp.27486.

[157] S. L. Edwards, J. Beesley, J. D. French, and A. M. Dunning, “Beyond GWASs: Illuminating the Dark Road from Association to Function,” *Am J Hum Genetics*, vol. 93, no. 5, pp. 779–797, 2013, doi: 10.1016/j.ajhg.2013.10.012.

[158] C. Perez-Cervantes *et al.*, “Enhancer transcription identifies cis-regulatory elements for photoreceptor cell types,” *Development*, vol. 147, no. 3, p. dev184432, 2020, doi: 10.1242/dev.184432.

[159] X. H. Yang *et al.*, “Transcription-factor-dependent enhancer transcription defines a gene regulatory network for cardiac rhythm.,” *Elife*, vol. 6, p. 1765, Dec. 2017, doi: 10.7554/elife.31683.

[160] M. T. Y. Lam, W. Li, M. G. Rosenfeld, and C. K. Glass, “Enhancer RNAs and regulated transcriptional programs,” *Trends Biochem Sci*, vol. 39, no. 4, pp. 170–182, 2014, doi: 10.1016/j.tibs.2014.02.007.

[161] D. Wang *et al.*, “Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA,” *Nature*, vol. 474, no. 7351, pp. 390–394, 2011, doi: 10.1038/nature10006.

[162] O. Mikhaylichenko *et al.*, “The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription.,” *Genes & Development*, vol. 32, no. 1, pp. 42–57, Jan. 2018, doi: 10.1101/gad.308619.117.

[163] T.-K. Kim *et al.*, “Widespread transcription at neuronal activity-regulated enhancers,” *Nature*, vol. 465, no. 7295, pp. 182–187, 2010, doi: 10.1038/nature09033.

[164] H. Wu *et al.*, “Tissue-Specific RNA Expression Marks Distant-Acting Developmental Enhancers,” *Plos Genet*, vol. 10, no. 9, p. e1004610, 2014, doi: 10.1371/journal.pgen.1004610.

[165] M. S. Werner *et al.*, “Chromatin-enriched lncRNAs can act as cell-type specific activators of proximal gene transcription,” *Nat Struct Mol Biol*, vol. 24, no. 7, pp. 596–603, 2017, doi: 10.1038/nsmb.3424.

[166] X. Qu *et al.*, “MIAT Is a Pro-fibrotic Long Non-coding RNA Governing Cardiac Fibrosis in Post-infarct Myocardium,” *Sci Rep-uk*, vol. 7, no. 1, p. 42657, 2017, doi: 10.1038/srep42657.

[167] S. Huang *et al.*, “Long noncoding RNA MALAT1 mediates cardiac fibrosis in experimental postinfarct myocardium mice model,” *J Cell Physiol*, vol. 234, no. 3, pp. 2997–3006, 2019, doi: 10.1002/jcp.27117.

[168] R. Micheletti *et al.*, “The long noncoding RNA Wisper controls cardiac fibrosis and remodeling.,” *Sci Transl Med*, vol. 9, no. 395, p. eaai9118, Jun. 2017, doi: 10.1126/scitranslmed.aai9118.

[169] M.-T. Piccoli *et al.*, “Inhibition of the Cardiac Fibroblast–Enriched lncRNA Meg3 Prevents Cardiac Fibrosis and Diastolic Dysfunction,” *Circ Res*, vol. 121, no. 5, pp. 575–583, 2017, doi: 10.1161/circresaha.117.310624.

[170] J. Viereck *et al.*, “Long noncoding RNA Chast promotes cardiac remodeling,” *Sci Transl Med*, vol. 8, no. 326, p. 326ra22, 2016, doi: 10.1126/scitranslmed.aaf1475.

[171] J.-F. Chen *et al.*, “The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation,” *Nat Genet*, vol. 38, no. 2, pp. 228–233, 2006, doi: 10.1038/ng1725.

[172] Y. Zhao *et al.*, “Dysregulation of Cardiogenesis, Cardiac Conduction, and Cell Cycle in Mice Lacking miRNA-1-2,” *Cell*, vol. 129, no. 2, pp. 303–317, 2007, doi: 10.1016/j.cell.2007.03.030.

[173] Y. Zhao, E. Samal, and D. Srivastava, “Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis,” *Nature*, vol. 436, no. 7048, pp. 214–220, 2005, doi: 10.1038/nature03817.

[174] B. Yang *et al.*, “The muscle-specific microRNA miR-1 regulates cardiac arrhythmogenic potential by targeting GJA1 and KCNJ2,” *Nat Med*, vol. 13, no. 4, pp. 486–491, 2007, doi: 10.1038/nm1569.

[175] J. Ai *et al.*, “Circulating microRNA-1 as a potential novel biomarker for acute myocardial infarction,” *Biochem Bioph Res Co*, vol. 391, no. 1, pp. 73–77, 2010, doi: 10.1016/j.bbrc.2009.11.005.

[176] U. Grabmaier *et al.*, “Diagnostic and prognostic value of miR-1 and miR-29b on adverse ventricular remodeling after acute myocardial infarction – The SITAGRAMI-miR analysis,” *Int J Cardiol*, vol. 244, pp. 30–36, 2017, doi: 10.1016/j.ijcard.2017.06.054.

[177] S. Zhou *et al.*, “miRNAS in cardiovascular diseases: potential biomarkers, therapeutic targets and challenges,” *Acta Pharmacol Sin*, vol. 39, no. 7, pp. 1073–1084, 2018, doi: 10.1038/aps.2018.30.

[178] E. Pinchi *et al.*, “miR-1, miR-499 and miR-208 are sensitive markers to diagnose sudden death due to early acute myocardial infarction,” *J Cell Mol Med*, vol. 23, no. 9, pp. 6005–6016, 2019, doi: 10.1111/jcmm.14463.

[179] F. Wang *et al.*, “miR-10a and miR-10b target the 3’-untranslated region of TBX5 to repress its expression.,” *Pediatr Cardiol*, vol. 35, no. 6, pp. 1072–1079, 2014, doi: 10.1007/s00246-014-0901-y.

[180] A. D. Hoffmann *et al.*, “Foxf genes integrate tbx5 and hedgehog pathways in the second heart field for cardiac septation.,” *PLoS Genet.*, vol. 10, no. 10, p. e1004604, 2014, doi: 10.1371/journal.pgen.1004604.

[181] J. Kim *et al.*, “Ablation of miR-10b Suppresses Oncogene-Induced Mammary Tumorigenesis and Metastasis and Reactivates Tumor-Suppressive Pathways,” *Cancer Res*, vol. 76, no. 21, pp. 6424–6435, 2016, doi: 10.1158/0008-5472.can-16-1571.

[182] D.-F. GUO *et al.*, “TBX5 loss-of-function mutation contributes to atrial fibrillation and atypical Holt-Oram syndrome,” *Mol Med Rep*, vol. 13, no. 5, pp. 4349–4356, 2016, doi: 10.3892/mmr.2016.5043.

[183] B. C. Bernardo *et al.*, “Therapeutic inhibition of the miR-34 family attenuates pathological cardiac remodeling and improves heart function,” *Proc National Acad Sci*, vol. 109, no. 43, pp. 17615–17620, 2012, doi: 10.1073/pnas.1206432109.

[184] T. G. Hullinger *et al.*, “Inhibition of miR-15 Protects Against Cardiac Ischemic Injury,” *Circ Res*, vol. 110, no. 1, pp. 71–81, 2012, doi: 10.1161/circresaha.111.244442.

[185] B. L. Murphy *et al.*, “Silencing of the miR-17~92 Cluster Family Inhibits Medulloblastoma Progression,” *Cancer Res*, vol. 73, no. 23, pp. 7068–7078, 2013, doi: 10.1158/0008-5472.can-13-0927.

[186] B. C. Bernardo *et al.*, "Inhibition of miR-154 Protects Against Cardiac Dysfunction and Fibrosis in a Mouse Model of Pressure Overload," *Sci Rep-uk*, vol. 6, no. 1, p. 22442, 2016, doi: 10.1038/srep22442.

[187] Q. Wang *et al.*, "miR-154-5p Functions as an Important Regulator of Angiotensin II-Mediated Heart Remodeling," *Oxid Med Cell Longev*, vol. 2019, p. 8768164, 2019, doi: 10.1155/2019/8768164.

[188] L. Ma *et al.*, "Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model," *Nat Biotechnol*, vol. 28, no. 4, pp. 341–347, 2010, doi: 10.1038/nbt.1618.

[189] L. Ma, J. Teruya-Feldstein, and R. A. Weinberg, "Tumour invasion and metastasis initiated by microRNA-10b in breast cancer," *Nature*, vol. 449, no. 7163, pp. 682–688, 2007, doi: 10.1038/nature06174.

[190] J. Zhang, L. Wu, Z. Li, and G. Fu, "miR-1231 exacerbates arrhythmia by targeting calciumchannel gene CACNA2D2 in myocardial infarction.," *Am J Transl Res*, vol. 9, no. 4, pp. 1822–1833, 2016.

[191] Surina, R. A. Fontanella, L. Scisciola, R. Marfella, G. Paolisso, and M. Barbieri, "miR-21 in Human Cardiomyopathies," *Frontiers Cardiovasc Medicine*, vol. 8, p. 767064, 2021, doi: 10.3389/fcvm.2021.767064.

[192] M. H. Kim, S. S. Johnston, B.-C. Chu, M. R. Dalal, and K. L. Schulman, "Estimation of Total Incremental Health Care Costs in Patients With Atrial Fibrillation in the United States," *Circulation Cardiovasc Qual Outcomes*, vol. 4, no. 3, pp. 313–320, 2011, doi: 10.1161/circoutcomes.110.958165.

[193] D. Schüttler *et al.*, "Animal Models of Atrial Fibrillation," *Circ Res*, vol. 127, no. 1, pp. 91–110, 2020, doi: 10.1161/circresaha.120.316366.

[194] H. A. Rockman *et al.*, "Segregation of atrial-specific and inducible expression of an atrial natriuretic factor transgene in an in vivo murine model of cardiac hypertrophy.," *Proc National Acad Sci*, vol. 88, no. 18, pp. 8277–8281, 1991, doi: 10.1073/pnas.88.18.8277.

[195] T. P. Nguyen, Z. Qu, and J. N. Weiss, "Cardiac fibrosis and arrhythmogenesis: The road to repair is paved with perils," *J Mol Cell Cardiol*, vol. 70, pp. 83–91, 2014, doi: 10.1016/j.yjmcc.2013.10.018.

[196] S. Anders, P. T. Pyl, and W. Huber, "HTSeq—a Python framework to work with high-throughput sequencing data," *Bioinformatics*, vol. 31, no. 2, pp. 166–169, 2015, doi: 10.1093/bioinformatics/btu638.

[197] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome Biol*, vol. 15, no. 12, p. 550, 2014, doi: 10.1186/s13059-014-0550-8.

[198] M. Zafeiriou, C. Noack, and L. Zelarayan, “Isolation and Primary Culture of Adult Mouse Cardiac Fibroblasts,” *Bio-protocol*, vol. 6, no. 13, 2016, doi: 10.21769/bioprotoc.1860.

[199] A. R. Ardakany, H. T. Gezer, S. Lonardi, and F. Ay, “Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation,” *Genome Biol*, vol. 21, no. 1, p. 256, 2020, doi: 10.1186/s13059-020-02167-0.

[200] Z. Arany, M. Novikov, S. Chin, Y. Ma, A. Rosenzweig, and B. M. Spiegelman, “Transverse aortic constriction leads to accelerated heart failure in mice lacking PPAR-γ coactivator 1α,” *Proc National Acad Sci*, vol. 103, no. 26, pp. 10086–10091, 2006, doi: 10.1073/pnas.0603615103.

[201] X. Zhang *et al.*, “Identification of Transcriptional Variation in Aortic Remodeling Using a Murine Transverse Aortic Constriction (TAC) Model,” *Frontiers Cardiovasc Medicine*, vol. 7, p. 581362, 2020, doi: 10.3389/fcvm.2020.581362.

[202] T. Doenst, T. D. Nguyen, and E. D. Abel, “Cardiac Metabolism in Heart Failure,” *Circ Res*, vol. 113, no. 6, pp. 709–724, 2013, doi: 10.1161/circresaha.113.300376.

[203] G. D. Lopaschuk, Q. G. Karwi, R. Tian, A. R. Wende, and E. D. Abel, “Cardiac Energy Metabolism in Heart Failure,” *Circ Res*, vol. 128, no. 10, pp. 1487–1513, 2021, doi: 10.1161/circresaha.121.318241.

[204] W. Li, D. Notani, and M. G. Rosenfeld, “Enhancers as non-coding RNA transcription units: recent insights and future perspectives,” *Nat Rev Genet*, vol. 17, no. 4, pp. 207–223, 2016, doi: 10.1038/nrg.2016.4.

[205] A. Mo *et al.*, “Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain,” *Neuron*, vol. 86, no. 6, pp. 1369–1384, 2015, doi: 10.1016/j.neuron.2015.05.018.

[206] S. K. Patel, J. Ramchand, V. Crocitti, and L. M. Burrell, “Kruppel-Like Factor 15 Is Critical for the Development of Left Ventricular Hypertrophy,” *Int J Mol Sci*, vol. 19, no. 5, p. 1303, 2018, doi: 10.3390/ijms19051303.

[207] S. Fisch *et al.*, “Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy,” *Proc National Acad Sci*, vol. 104, no. 17, pp. 7074–7079, 2007, doi: 10.1073/pnas.0701981104.

[208] Y. Yu *et al.*, “KLF15 Is an Essential Negative Regulatory Factor for the Cardiac Remodeling Response to Pressure Overload,” *Cardiology*, vol. 130, no. 3, pp. 143–152, 2015, doi: 10.1159/000369382.

[209] B. Wang *et al.*, “The Kruppel-like factor KLF15 inhibits connective tissue growth factor (CTGF) expression in cardiac fibroblasts,” *J Mol Cell Cardiol*, vol. 45, no. 2, pp. 193–197, 2008, doi: 10.1016/j.yjmcc.2008.05.005.

[210] L. E. Montefiori *et al.*, “A promoter interaction map for cardiovascular disease genetics,” *Elife*, vol. 7, p. e35788, 2018, doi: 10.7554/elife.35788.

[211] G. M. Scharf *et al.*, “Inactivation of Sox9 in fibroblasts reduces cardiac fibrosis and inflammation.,” *JCI Insight*, vol. 5, no. 15, p. e6, Jul. 2019, doi: 10.1172/jci.insight.126721.

[212] G. P. A. Lacraz *et al.*, “Tomo-Seq Identifies SOX9 as a Key Regulator of Cardiac Fibrosis During Ischemic Injury.,” *Circulation*, vol. 136, no. 15, pp. 1396–1409, Oct. 2017, doi: 10.1161/circulationaha.117.027832.

[213] P. R. Gajjala *et al.*, “Dysregulated overexpression of Sox9 induces fibroblast activation in pulmonary fibrosis,” *Jci Insight*, vol. 6, no. 20, p. e152503, 2021, doi: 10.1172/jci.insight.152503.

[214] X. Meng, D. J. Nikolic-Paterson, and H. Y. Lan, “TGF-β: the master regulator of fibrosis,” *Nat Rev Nephrol*, vol. 12, no. 6, pp. 325–338, 2016, doi: 10.1038/nrneph.2016.48.

[215] B. Schmierer and C. S. Hill, “TGFβ–SMAD signal transduction: molecular specificity and functional flexibility,” *Nat Rev Mol Cell Bio*, vol. 8, no. 12, pp. 970–982, 2007, doi: 10.1038/nrm2297.

[216] A. Acharya *et al.*, “The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors,” *Development*, vol. 139, no. 12, pp. 2139–2149, 2012, doi: 10.1242/dev.079970.

[217] O. Sazonova *et al.*, “Characterization of TCF21 Downstream Target Regions Identifies a Transcriptional Network Linking Multiple Independent Coronary Artery Disease Loci,” *Plos Genet*, vol. 11, no. 5, p. e1005202, 2015, doi: 10.1371/journal.pgen.1005202.

[218] A. Alvarez-Franco *et al.*, “Transcriptome and proteome mapping in the sheep atria reveal molecular featurets of atrial fibrillation progression,” *Cardiovasc Res*, vol. 117, no. 7, pp. 1760–1775, 2020, doi: 10.1093/cvr/cvaa307.

[219] B. S. Mankoo *et al.*, “The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites,” *Development*, vol. 130, no. 19, pp. 4655–4664, 2003, doi: 10.1242/dev.00687.

[220] A. Riddell *et al.*, “RUNX1: an emerging therapeutic target for cardiovascular disease,” *Cardiovasc Res*, vol. 116, no. 8, pp. 1410–1423, 2020, doi: 10.1093/cvr/cvaa034.

[221] C. S. McCarroll *et al.*, “Runx1 Deficiency Protects Against Adverse Cardiac Remodeling After Myocardial Infarction,” *Circulation*, vol. 137, no. 1, pp. 57–70, 2018, doi: 10.1161/circulationaha.117.028911.

[222] S. Gattenlöhner, C. Waller, G. Ertl, B.-D. Bültmann, H.-K. Müller-Hermelink, and A. Marx, “NCAM(CD56) and RUNX1(AML1) Are Up-Regulated in Human Ischemic

Cardiomyopathy and a Rat Model of Chronic Cardiac Ischemia," *Am J Pathology*, vol. 163, no. 3, pp. 1081–1090, 2003, doi: 10.1016/s0002-9440(10)63467-0.

[223] T. P. Martin *et al.*, "Inhibiting Runx1 protects heart function after myocardial infarction," *Biorxiv*, p. 2022.02.17.480749, 2022, doi: 10.1101/2022.02.17.480749.

[224] L. Zhang *et al.*, "Forkhead Box S1 mediates epithelial-mesenchymal transition through the Wnt/β-catenin signaling pathway to regulate colorectal cancer progression," *J Transl Med*, vol. 20, no. 1, p. 327, 2022, doi: 10.1186/s12967-022-03525-1.

[225] T. A. Blenkinsop *et al.*, "FOXS1 is a Master Regulator of Pathological Epithelial to Mesenchymal Transition in Human Epithelia," *Biorxiv*, p. 154369, 2017, doi: 10.1101/154369.

[226] S. Wang *et al.*, "FOXS1 is regulated by GLI1 and miR-125a-5p and promotes cell proliferation and EMT in gastric cancer," *Sci Rep-uk*, vol. 9, no. 1, p. 5281, 2019, doi: 10.1038/s41598-019-41717-w.

[227] K. Bévant *et al.*, "TGFβ-induced FOXS1 controls epithelial–mesenchymal transition and predicts a poor prognosis in liver cancer," *Hepatology Commun*, vol. 6, no. 5, pp. 1157–1171, 2021, doi: 10.1002/hep4.1866.

[228] T. M. Consortium *et al.*, "Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris.," *Nature*, vol. 562, no. 7727, pp. 367–372, 2018, doi: 10.1038/s41586-018-0590-4.

[229] S. M. Haldar *et al.*, "Klf15 Deficiency Is a Molecular Link Between Heart Failure and Aortic Aneurysm Formation," *Sci Transl Med*, vol. 2, no. 26, p. 26ra26, 2010, doi: 10.1126/scitranslmed.3000502.

[230] J. J. Leenders *et al.*, "Regulation of Cardiac Gene Expression by KLF15, a Repressor of Myocardin Activity*," *J Biol Chem*, vol. 285, no. 35, pp. 27449–27456, 2010, doi: 10.1074/jbc.m110.107292.

[231] H. Qin *et al.*, "SOX9 in prostate cancer is upregulated by cancer-associated fibroblasts to promote tumor progression through HGF/c-Met-FRA1 signaling," *Febs J*, vol. 288, no. 18, pp. 5406–5429, 2021, doi: 10.1111/febs.15816.

[232] H. Li *et al.*, "TGF-β-mediated upregulation of Sox9 in fibroblast promotes renal fibrosis," *Biochimica Et Biophysica Acta Bba - Mol Basis Dis*, vol. 1864, no. 2, pp. 520–532, 2018, doi: 10.1016/j.bbadis.2017.11.011.

[233] C. P. D. Wheeler-Jones, "Cell signalling in the cardiovascular system: an overview," *Heart*, vol. 91, no. 10, p. 1366, 2005, doi: 10.1136/hrt.2005.072280.

[234] F. G. Ottaviano and K. O. Yee, "Communication Signals Between Cardiac Fibroblasts and Cardiac Myocytes," *J Cardiovasc Pharm*, vol. 57, no. 5, pp. 513–521, 2011, doi: 10.1097/fjc.0b013e31821209ee.

[235] C. Hall, K. Gehmlich, C. Denning, and D. Pavlovic, “Complex Relationship Between Cardiac Fibroblasts and Cardiomyocytes in Health and Disease,” *J Am Hear Assoc Cardiovasc Cerebrovasc Dis*, vol. 10, no. 5, p. e019338, 2021, doi: 10.1161/jaha.120.019338.

[236] V. Dusi, A. Ghidoni, A. Ravera, G. M. D. Ferrari, and L. Calvillo, “Chemokines and Heart Disease: A Network Connecting Cardiovascular Biology to Immune and Autonomic Nervous Systems,” *Mediat Inflamm*, vol. 2016, p. 5902947, 2016, doi: 10.1155/2016/5902947.

[237] A. Conesa, M. J. Nueda, A. Ferrer, and M. Talon, “maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments,” *Bioinformatics*, vol. 22, no. 9, pp. 1096–1102, 2006, doi: 10.1093/bioinformatics/btl056.

[238] M. J. Nueda, S. Tarazona, and A. Conesa, “Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series,” *Bioinformatics*, vol. 30, no. 18, pp. 2598–2602, 2014, doi: 10.1093/bioinformatics/btu333.

[239] X. Shao, J. Liao, C. Li, X. Lu, J. Cheng, and X. Fan, “CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice.,” *Brief Bioinform*, vol. 22, no. 4, 2020, doi: 10.1093/bib/bbaa269.

[240] J. A. Fontes, N. R. Rose, and D. Čiháková, “The varying faces of IL-6: From cardiac protection to cardiac failure,” *Cytokine*, vol. 74, no. 1, pp. 62–68, 2015, doi: 10.1016/j.cyto.2014.12.024.

[241] F. Chen *et al.*, “Hop Is an Unusual Homeobox Gene that Modulates Cardiac Development,” *Cell*, vol. 110, no. 6, pp. 713–723, 2002, doi: 10.1016/s0092-8674(02)00932-7.

[242] C. H. Shin *et al.*, “Modulation of Cardiac Growth and Development by HOP, an Unusual Homeodomain Protein,” *Cell*, vol. 110, no. 6, pp. 725–735, 2002, doi: 10.1016/s0092-8674(02)00933-9.

[243] H. Kook *et al.*, “Cardiac hypertrophy and histone deacetylase–dependent transcriptional repression mediated by the atypical homeodomain protein Hop,” *J Clin Invest*, vol. 112, no. 6, pp. 863–871, 2003, doi: 10.1172/jci19137.

[244] C. M. Trivedi, T. P. Cappola, K. B. Margulies, and J. A. Epstein, “Homeodomain Only Protein X is down-regulated in human heart failure,” *J Mol Cell Cardiol*, vol. 50, no. 6, pp. 1056–1058, 2011, doi: 10.1016/j.yjmcc.2011.02.015.

[245] X. Yu *et al.*, “Reduced expression of BMP3 contributes to the development of pulmonary fibrosis and predicts the unfavorable prognosis in IIP patients,” *Oncotarget*, vol. 8, no. 46, pp. 80531–80544, 2017, doi: 10.18632/oncotarget.20083.

[246] B. Song *et al.*, “Inhibition of BMP3 increases the inflammatory response of fibroblast-like synoviocytes in rheumatoid arthritis,” *Aging Albany Ny*, vol. 12, no. 12, pp. 12305–12323, 2020, doi: 10.18632/aging.103422.

[247] Z. Wang *et al.*, “The long noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy,” *Nat Med*, vol. 22, no. 10, pp. 1131–1139, 2016, doi: 10.1038/nm.4179.