UNIVERSITY OF CALIFORNIA

Los Angeles

RNA Splicing Regulation in Cardiac Development and Disease

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

Chen Gao

2014

ABSTRACT OF THE DISSERTATION

RNA Splicing Regulation In Cardiac Development and Disease

By

Chen Gao

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2014

Professor Yibin Wang, Chair

During cardiac development and pathological remodeling, there is a transcriptome maturation and remodeling event well established at transcription level. However, with high-throughput sequencing technology, we are able to obtain a more comprehensive understanding of the real transcriptome complexity at single base resolution. In order to understand the cardiac transcriptome complexity and dynamics during normal and disease conditions, we performed deep RNA-Sequencing on pressured overload induced mouse failing hearts and compared with sham operated control hearts. From this study, we have identified a significant number of genes undergo alternative splicing during heart disease. We have also provided evidence there is a large number of previously un-annotated novel splicing variants, lncRNA and novel transcript clusters, some of these could have potential impact on cardiac disease.

From the sequencing analysis, we chose to carry out detailed characterization of a novel cardiac specific splicing variant in PKCα. Both biochemistry and cell studies suggested this novel splicing variant has significant higher auto-phosphorylation level at baseline but has different activation profile responding to hypertrophic stimuli. This is potentially due to this novel PKCα splicing variant has unique interacting partner and downstream target. We further demonstrated that, the alternative splicing of this novel PKCα variant is, at least partially regulated by RBFox1.

RNA splicing contributes significantly to total transcriptome complexity but its functional role and regulation in cardiac development and diseases remain poorly understood. Based on total transcriptome analysis, we identified a significant number of alternative RNA splicing events in mouse failing hearts that resembled the pattern in fetal hearts. A muscle specific isoform of an RNA splicing regulator RBFox1 (A2BP1) is induced during cardiac development. Inactivation of zRBFox1 gene in zebrafish led to lethal phenotype associated with impaired cardiac function. RBFox1 regulates alternative splicing of transcription factor MEF2s, producing splicing variants with distinct transcriptional activities and different impact on cardiac development. RBFox1 expression is diminished in mouse and human failing hearts. Restoring RBFox1 expression significantly attenuates hypertrophy and heart failure induced by pressure-overload in mice. Therefore, RBFox1-MEF2 represents a previously uncharacterized regulatory circuit in cardiac transcriptional network with important impact on both cardiac development and diseases.

The dissertation of Chen Gao is approved by

Douglas L. Black

Xinshu (Grace) Xiao

Yibin Wang, Committee Chair

University of California, Los Angeles

2014

Dedication

To everyone who walks me through

the rain

J'aime la vie

TABLE OF CONTENTS AND LISTS

LISTS OF TABLES

LISTS OF FIGURES

ACKNOWLEDGMENTS

I am really grateful to meet my mentor: Dr. Yibin Wang, who has been giving me tremendous support since I was a college student. Yibin is the one who always provides insights and constructive suggestions to all my projects, but his role is far beyond a scientific mentor. Yibin spent enormous amount of time teaching me how to communicate my scientific story; how to collaborate with my colleague; how to overcome difficulties in my project; how to embrace challenges in my career, and in my life as well. His enthusiasm in educating student; his courage to take on novel and challenging questions; his willingness to help and support others inspired me during my graduate study and will continue inspiring me in my future career and life.

I also wish to express my sincere appreciation to all my committee members, including Dr. Jaunian Chen, Dr. Xinshu Xiao, Dr. Douglas L Black and Dr. James Weiss, who have given me almost endless support, resource and advice during my four years as their student. I would like to thank Dr. Jaunian Chen and everyone in her lab teaching me all the zebrafish experiments and accommodating me with the resource. I would like to thank Dr. Xinshu Xiao and Jae-Hyung Lee for all the bioinformatics analysis, without their help, the splicing project would never happen. I would like to thank Dr. Douglas L Black for educating me all the

alternative splicing knowledge and experimental tools, and providing me all the necessary resources for RBFox1 project. I would also like to thank Dr. James Weiss, who always gives me insights on cardiac physiology and pathology.

None of my thesis projects can be accomplished without the technical support from my lab members. I would like to thank Dr. Vincent Ren who supported me with TAC surgery; I would like to thank Haiying Pu and Jing Gao for their mice breeding and genotyping work; I would also like to thank my undergraduate students who have helped me with all the work I don't have time to do.

The project is supported by UCLA Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research Center Training Grant and UCLA Dissertation Year Fellowship.

And most importantly, I would like to thank my parents. For all the tough times, they are there for me, supporting me and comforting me, as always.

BIOGRAPHICAL SKETCH

USE ONLY FOR INDIVIDUAL PREDOCTORAL and POSTDOCTORAL FELLOWSHIPS. DO NOT EXCEED FOUR PAGES.

NAME OF FELLOWSHIP APPLICANT	POSITION TITLE
CHEN GAO, BS	Graduate Student
eRA COMMONS USER NAME (credential, e.g., agency login) chengao	

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Nankai University,Tianjin,China	BS	2009	Biology Technology

Please refer to the application instructions in order to complete sections A, B, C, and D of the Biographical Sketch.

A. Personal Statement

B. Positions and Honors

ACTIVITY/OCCUPATION	BEGINNING DATE (mm/yy)	ENDING DATE (mm/yy)	FIELD	INSTITUTION/COMPANY	SUPERVISOR/ EMPLOYER
College Student	2005	2009	Biology	Nankai Univerisity, China	Jin-tang Dong
Graduate Student	2009	now	Molecular Biology	Univeristy of California,Los Angeles	Yibin Wang

Academic and Professional Honors

2013 International Society of Heart Research (ISHR) Richard J. Bing Award for Young Investigators

2013 Eli& Edythe Broad Center of Regenerative Medicine& Stem Cell Research Center Training Grant

2013 UCLA Dissertation Year Fellowship

2012 Eli& Edythe Broad Center of Regenerative Medicine& Stem Cell Research Center Training Grant

2011 AHA BCVS New Investigator Travel Award

2009 Nankai Excellent Undergraduate Thesis Award

2008 Nankai Third Prize Scholarship for excellent academic performance & honored excellent student

2007 Xinan Union University Scholarship

2006 Nankai Secondary Prize Scholarship for excellent academic performance & honored excellent student

C. Publications(peer-reviewed)

1. RBFox1 Mediated Alternative RNA Splicing Regulates Postnatal Development and Function in Heart, **Chen Gao**, Shuxun Ren, Jae-Hyung Lee, Haodong Chen, Jared Churko, Joseph Wu, Thomas Vondriska, Preeti Ahuja, Xinshu Xiao, Jau-Nian Chen, Yibin Wang (manuscript submitted)
2. Nuclear phosphatase PPM1G in cellular survival and neural development, William H. Foster, Adam Langenbacher, **Chen Gao**, Jaunian Chen, Yibin Wang; *Dev Dyn*, 2013 July 26
3. Global impact of RNA splicing on transcriptome remodeling in the heart, **Chen Gao**, Yibin Wang *J Zhejiang Univ Sci B*, 2012 Aug; 13(8); 603-8
4. Tissue-specific and nutrient regulation of the branched-chain α-keto acid dehydrogenase phosphatase, protein phosphatase 2Cm (PP2Cm), Meiyi Zhou, **Chen Gao**, Gang Lu, Yibin Wang and Haipeng Sun *J Biol Chem*, 2012 Jul 6; 287(28);23397-406.
5. Analysis of transcriptome complexity through RNA sequencing in normal and failing murine hearts, Jae-hyung Lee, **Chen Gao**, Guangdun Peng,Christopher Greer,Shuxun Ren, Yibin Wang, Xinshu Xiao *Circulation Research* 2011 Dec 9; 109(12): 1332-41
6. Expression, purification and characterization of a thermophilic neutral protease from *Bacillus stearothermophilus* in *Bacillus subtilis* ZHANG Min, ZHAO Cong, DU Lianxiang, LU Fuping& **GAO Chen** *Sci China Ser C-Life Sci* Jan.2008
7. Characterization of mutagenised acid-resistant alpha-amylase expressed in *Bacillus subtilis* WB600 Yihan Liu, Fuping Lu, Xiangbin Yin, Yi Wang, **Chen Gao** *Appl Microbiol Biotechnol* Dec.2007
8. High-level production and characterization of the recombinant thermostable lipase of *Geobacillus thermoleovorans* in *Pichia methanolica* Wei Qi, Haikuan Wang, Ruijuan Liu, **Chen Gao**, Fuping Lu *Annals of Microbiology* Jan 2008

D. Conference and Presentation

1. "Global RNA Splicing and Regulation in Cardiac Maturation and Diseases", poster presentation, Basic Cardiovascular Sciences, 2013 Scientific Sessions, Las Vegas, Nevada
2. "RBFox1 Mediated Alternative RNA Splicing of MEF2 Transcripts Regulates Postnatal Development and Function in Heart", oral presentation, 2013 International Society of Heart Research World Congress, San Diego, CA
3. "Global RNA Splicing and Regulation in Fetal and Diseased Hearts", poster presentation, American Heart Association Late-breaking Basic Science, 2012 American Heart Association Scientific Sessions, Los Angeles, CA
4. "Fetal-like RNA Splicing Remodeling in Failing Heart Revealed by Total Transcriptome Analysis", poster presentation, Basic Cardiovascular Sciences, 2012 Scientific Sessions, New Orleans
5. "Identification of a Novel PKCα Isoform in Heart Transcriptome via RNA-Deep Sequencing", poster presentation, Basic Cardiovascular Sciences, 2012 Scientific Sessions, New Orleans
6. "Global Transcriptome Remodeling in Heart Failure--New Insights Revealed from RNA-Seq", poster presentation, Basic Cardiovascular Sciences, 2011 Scientific Sessions, New Orleaans

Chapter One

Introduction

Cardiac Development and Pathological Remodeling

Early cardiac development contains a series of precisely orchestrated molecular and morphogenetic events that ultimately lead to a functional mature four-chamber heart(Deepak Srivastava 2000, Olson 2002, Joanna Kobylinska 2013, Rana, Christoffels et al. 2013). Using genetic loss-of-function studies in model system, we are able to understand the major molecular events that are critical for early cardiac development.

Previous study has demonstrated that tissue specific transcription factors play critical roles controlling the cardiac cell fate decision differentiated from mesodermal stem cells. During this transition, Nkx2.5 has been identified as the earliest molecular marker that defines the cardiomyocytes lineage(Harvey 1996), which directly leads to the activation of MEF2 genes(Gajewski, Kim et al. 1998). Another transcription factor that should be mentioned here is the zinc-finger transcription factor of the GATA family(Durocher, Charron et al. 1997, Turbendian, Gordillo et al. 2013), together with MEF2 family transcription factors, they activate cardiac gene expression during early cardiomyocytes fate determination.

Earlier study has also identified key transcription factors regulating left and right ventricles development. The basic helix-loop-helix (bHLH) transcription factors dHAND1/HAND2 and MEF2 seem to be critical for both left and right ventricle development(Srivastava, Cserjesi et al. 1995, Lin, Schwarz et al. 1997, Srivastava 1997), the left and right ventricle also express chamber specific transcription

factors. While the Versican has been demonstrated to have a critical role during left ventricle development(Henderson and Copp 1998, Mjaatvedt, Yamamura et al. 1998), the TBX proteins, on the other hand, seem to be more important regulating right ventricle development(Greulich, Rudat et al. 2011).

In addition to transcription factor expression and regulation, the fetal heart is also different from the mature adult heart in other aspect. The cardiac metabolism before birth relies on using carbohydrate as energy provision, however, the oxidation of fatty acids turned into the predominant form of metabolism in adult heart(Taegtmeyer, Sen et al. 2010).

Interestingly, during cardiac stress, there is a remodeling event occurs at metabolic level. A variety of pathological conditions, including cardiac hypertrophy and heart failure can trigger the stressed heart changing the major energy substrate back into glucose(Taegtmeyer, Sen et al. 2010).

In addition to the change of metabolism, at gene expression level, the stressed heart also returns to " fetal gene program"(Barry, Davidson et al. 2008, Kuwahara, Nishikimi et al. 2012). During the fetal gene reprogramming, genes are abundantly expressed in fetal ventricles are re-induced, including: atrial and brain natriuretic peptide; fetal isoforms of contractile proteins (skeletal a-actin and b-myosin heavy chain), fetal type cardiac ion channels (hyperpolarization-activated cyclic nucleotide-gated channel and T-type Calcium channel) as well as some smooth muscle genes.

Current Progress in Cardiac Transcriptome Analysis

During the past decade, microarray and proteomic studies have provided a wealth of knowledge on cardiac transcriptome complexity during cardiac development and pathological remodeling.

Firstly, we have demonstrated that cardiac transcriptome maturation and pathological remodeling is regulated at gene expression level by key transcription factors, including the Nkx2.5, MEF2,GATA and TBX proteins that are mentioned in the previous section.

Secondly, proteomics study has also revealed the functional impact of signaling pathway during cardiac development and disease. One of the most well established pathway-MAPK (mitogen-activated protein kinase) has been proved to play a critical role mediating cardiomyocytes hypertrophy response both in vitro cultured cardiomyocytes and in intact hearts. This pathway is regulated through phosphorylation of the key components, including p38, ERK and Akt(Sharma GD 2002, Kilic, Velic et al. 2005).

Histone modification is another important pathway mediating cardiac hypertrophy. Overexpression of histone acetyltransferases-CBP/p300 is sufficient to induced cardiac hypertrophy and remodeling in transgenic mouse models(Gusterson, Jazrawi et al. 2003, Yanazume, Hasegawa et al. 2003). In addition to histone acetyltransfereases, there are also more than one dozen individual HDACs mediating cardiomyocytes hypertrophy response. By interacting with MEF2 transcription factors, the HDACs can both activate or suppress cardiomyocytes hypertrophy response(Lu, McKinsey et al. 2000, Zhang, McKinsey et al. 2002, Chang, McKinsey et al. 2004).

Lastly, the microRNA has been studied extensively as a novel regulator in cardiac hypertrophy and heart failure. With miRNA microarray approach, a significant number of miRNAs have been identified to associate with cardiac stress(Orenes-Piñero, Montoro-García et al. 2013). Among them, miR-22 has been suggested to induce hypertrophy response in vitro cultured cardiomyocytes, and loss of miR22 also repressed Calcineurin-induced cardiac hypertrophy(Huang, Chen et al. 2013). Another good example is miR-212 and miR-132. By directly targeting the anti-hypertrophic and pro-autophagic FoxO3 transcription factors, these two miR further activate hypertrophic Calcineurin/NFAT signaling pathway in cardiomyocytes(Ucar, Gupta et al. 2012).

In summary, with gene expression microarray, miR-microarray and proteomics studies, the complexity of cardiac transcriptome has been expanded significantly. During cardiac development and pathological remodeling, the histone modification enzymes regulate chromatin status; the key transcription factors regulate total gene expression level, the miR further regulates gene expression at post-transcriptional level and the complexity of transcriptome is further expanded by the stress signaling pathways that modify proteins at post-translational level. However, with the appearance of deep RNA-sequencing, our understanding of transcriptome has reached single exon resolution.

Expanding Transcriptome Complexity by Deep RNA Sequencing

For a long time, our understanding of transcriptome is limited by biased gene prediction and EST evidence. However, the development of deep RNA-Sequencing technology has, for the first time, revealed the complex landscape and

dynamics of the transcriptome at a much higher level of accuracy. In contrast to traditional sequencing technology and microarray study, deep RNA-Sequencing allows us to zoom into the transcriptome at single base-pair-resolution and de novo annotate transcripts(Birol 2009, Wang, Gerstein et al. 2009, Au, Jiang et al. 2010, Adamidi 2011). With these advantages, more and more studies have provided a more comprehensive understanding of transcriptome including both large and small RNAs, novel transcripts from unannotated genes and splicing isoforms(Martin and Wang 2011).

In order to obtain a more comprehensive understanding of the cardiac transcriptome complexity, especially the dynamics of cardiac transcriptome under normal and pathological conditions, we have performed deep RNA-Sequencing on pressure-overload induced early stage hypertrophy and end stage failing hearts and compared with sham operated control hearts(Lee, Gao et al. 2011). In addition to total gene expression analysis, we have also developed new bioinformatics tools that identified large amount of differentially expressed transcript isoforms; novel spliced exons, novel alternative terminal exons, novel transcript clusters as well as long noncoding RNA genes. Our study, together with a following study also using deep RNA-Sequencing technology(Hong Ki Song 2012), has significantly expanded our understanding of the total cardiac transcriptome complexity under different physiological and pathological stages.

PKCα-NE--- A Cardiac Specific Novel PKCα Splicing Variant

Among the discoveries we made during this deep sequencing effort, we are particularly interested in a novel exon in PKCα.

The protein kinase C (PKC) family is a critical regulator of cardiac signal transduction. Based on their activation mechanism, the PKC family members are divided into conventional PKC isozymes including PKCα, which respond to Calcium and lipid activation; while the novel and atypical isozymes are Calcium independent but can be activated by lipid(Eric Chruchill 2008, Steinberg 2008). Earlier studies have demonstrated a critical role of PKCα mediating cardiac function and cardiomyocytes contractility. In vitro study using human ventricular cardiomyocytes suggested PKCα translocation from cytosol to the contractile system plays an important role maintaining the contractile force of cardiomyocytes by phosphorylating its downstream target—cardiac Troponin I (TnI)(Molnár, Borbély et al. 2009), in vivo pharmacological inhibition of PKCα using ruboxistaurin has been demonstrated to have antagonizing effect on heart failure post myocardial infarction injury, potentially by regulating cardiac contractility, myocyte cellular contractility, Calcium transient and sarcoplasmic reticulum Calcium load(Hambleton, Hahn et al. 2006, Liu, Chen et al. 2009, Ladage, Tilemann et al. 2011). The regulatory mechanism of PKCα is also well established. The newly synthesized PKC has an open conformation that allows the PDK-1 to phosphorylate its priming phosphorylation site, in order to be further activated, the cPKC, including PKCα needs two additional phosphorylation events, one in turn motif and the other in hydrophobic motif. These autophosphorylation events are critical for PKCα activity by affecting the enzyme thermal stability, detergent solubility as well as protease/phosphatase susceptibility(Edwards and Newton 1997, Steinberg 2008).

Our study, however, has identified a novel regulatory mechanism of PKCα at alternative splicing level. Based on our deep RNA-Sequencing analysis, we have found a previously un-annotated novel exon of PKCα inserted right in front of the original protein turn motif, and will have a significant impact on the protein structure based on protein structure prediction. In vitro and in vivo studies have shown the insertion of novel exon in PKCα generated a significant higher level of autophosphorylation at turn motif. Surprisingly, this PKCα-NE (Novel Exon) also has very unique activation profile in cultured cardiomyocytes upon PMA (phorbol 12-myristate 13-acetate), Isoproterenol and Angiotensin II stimulation. Moreover, in contrast to PKCα-WT, the PKCα-NE failed to phosphorylate classic PKCα downsteam target-TnI upon Angiotensin II treatment. In order to determine the binding partner and downstream target of PKCα-NE, we performed immuno-precipitation study followed by mass spectrometry. Interestingly, comparing to PKCα-WT, the PKCα-NE has its unique interacting partners, including key components of protein translation machinery—eEF1A1. We have further demonstrated that eEF1A1 is indeed interacting with PKCα-NE based on immune-precipitation study in cultured cardiomyocytes; and can be potentially phosphorylated by PKCα-NE. Lastly, we investigated the regulatory mechanism of PKCα novel exon alternative splicing. Based on bioinformatics analysis, this highly conserved novel exon also shares conserved flanking cis-regulatory elements across different species. Within the conserved cis-regulatory elements, we have identified two putative RBFox1 binding motifs. Both in vivo and in vitro

minigene reporter analysis have provided evidence that RBFox1 indeed regulates the PKCα novel exon splicing directly.

In summary, our deep RNA-sequencing analysis has identified a previously un-annotated cardiac specific splicing event of PKCα. Functional characterization suggested the insertion of this novel exon could have a significant impact on the enzyme activation responding to hypertrophic stimuli; the insertion of the novel exon can also affect the enzyme activity towards downstream target, thus providing a novel regulatory mechanism for this well establish cardiac signal regulator.

Cardiac Genes Undergo Extensive Alternative Splicing

In addition to the novel splicing variant of PKCα, our deep RNA-Sequencing study has also identified a significant number of genes undergo alternative splicing during cardiac hypertrophy and failure comparing to sham operated hearts(Lee, Gao et al. 2011).

Alternative splicing plays an important role regulating cardiac gene expression. One extreme example of alternative splicing is Titin. This extremely large protein has been identified to have multiple splicing variants that are differentially expressed during cardiac development and disease(Wei Guo 2010). Recent study has further identified a splicing regulator-RBM20 that is responsible for regulating Titin alternative splicing(Guo, Schafer et al. 2012).

Another gene that also undergoes extensive alternative splicing is VEGF (Vascular Endothelial Growth Factor). This key regulator of angiogenesis contains eight exons. Alternative splicing of VEGF generates a variety of

isoforms that are different in both structure and function, leading to either pro-angiogenic activity or anti-angiogenic activity(Tischer, Gospodarowicz et al. 1989, Jingjing, Xue et al. 1999, Manetti, Guiducci et al. 2011). Further study has also identified the regulator—SC35 for VEGF alternative splicing. This SR protein splicing regulator, by interacting with transcription factor, regulates the ratio of pro-angiogenic and anti-angiogenic VEGF splicing variants expression, at least in p53 deficient tumor cells(Merdzhanova, Gout et al. 2010).

Cardiac channel proteins are also regulated at alternative splicing level during both cardiac transcriptome maturation and pathological remodeling. One example that has been studied extensively is SCN5A. This sodium channel protein showed abnormal splicing pattern under a variety of pathological conditions, including myotonic dystrophy type 1 (DM1) and heart failure-associated arrhythmia.(Wahbi, Algalarrondo et al. , Murphy, Moon-Grady et al. 2012, Jr 2013) The abnormal splicing pattern of SCN5A in human heart failure has been suggested to be regulated by LUC7L3 and RBM25(Gao and Dudley Jr 2013).

Our study, on the other hand, for the first time revealed a global alternative splicing events associated with both cardiac hypertrophy and heart failure. According to our deep RNA-Sequencing analysis in failing murine heart induced by pressure-overload, there are a total of 7811 isoforms expressing in heart being detected. We have also identified 1087 genes with significant isoform-specific expression changes, and a total of 720 isoforms in 475 genes are differentially expressed between normal and diseased hearts. We also showed more genes undergo alternative splicing during heart failure stage comparing to

hypertrophy(Lee, Gao et al. 2011). The scale of alternative splicing in heart under disease is also reviewed in our and others recent reviews(Chen Gao 2012, Lara-Pezzi, Gómez-Salinero et al. 2013, Zhang and Shaw 2013).

RBFox1 Mediated RNA Splicing Regulation During Cardiac Development and Disease

In my thesis study, we have identified a cardiac enriched splicing regulator—RBFox1 playing a critical role mediating cardiac transcriptome maturation and pathological remodeling. Bioinformatics analysis showed RBFox1 binding motif is highly enriched in the alternative splicing events associated with heart failure. The expression level of RBFox1 is dynamically regulated during both cardiac development and pathological remodeling at both mRNA and protein levels. As a conserved splicing regulator, RBFox1 plays a critical role in zebrafish cardiac development and function demonstrated in RBFox1 morphant embryos. Further, we have identified a downstream target—MEF2 family to be directly regulated by RBFox1. The conserved splicing event of MEF2 generated a mutually exclusive adult VS fetal splicing variant. RNA-Sequencing analysis combined with in vitro luciferase reporter assay suggested different MEF2 splicing variants have distinct transcription factor activities. Lastly, we have demonstrated that RBFox1 also has important role mediating cardiac hypertrophy response. Overexpression of RBFox1 is sufficient to attenuate hypertrophy response induced by PE treatment in vitro cultured cardiomyocytes; in vivo, cardiac specific overexpression of RBFox1 is also sufficient to preserve mice cardiac function post pressure-overload induced heart failure.

Thesis Project and Goals

This Chapter has provided introduction of my thesis projects, including the background of the heart failure and research tools that play important roles for my study. The following chapters will be as follow: Chapter Two: Global Transcriptome Analysis in Pressure-Overload Induced Mouse Failing Heart, including a new method paper published in Circulation Research in 2012, which summarized my major findings in transcriptome complexity in mouse normal and failing heart using RNA-Sequencing technique. Chapter Three: Functional Characterization of a Novel Exon in PKCα in Heart, gives an example of cardiac transcriptome complexity, where we have identified a previously un-annotated exon in PKCα, and demonstrated the insertion of this novel exon can have a major functional impact on the original protein; Chapter Four: Cardiac Genes Undergo Extensive Alternative Splicing, including a review that I wrote with Dr. Yibin Wang on our findings in RNA-Sequencing and research from other groups on how extensive cardiac genes undergo alternative splicing during normal and pathological states; providing a strong rationale for us to further dissecting the molecular mechanism regulating alternative splicing in heart. Chapter Five: RBFox1 Mediated Alternative RNA Splicing Regulates Development and Function in Heart. This is the major component of my thesis project, where we have identified a cardiac splicing regulator—RBFox1, by regulating the mutually exclusive alternative splicing of MEF2 family members, regulating both cardiac development and function. Chapter Six is conclusion and closing remarks, with discussion of future directions in my projects and therapeutic value of my study.

References:

Adamidi, C. (2011). "De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics." Genome Res. **21**: 1193-1200.

Au, K. F., H. Jiang, L. Lin, Y. Xing and W. H. Wong (2010). "Detection of splice junctions from paired-end RNA-seq data by SpliceMap." Nucleic Acids Res. **38**: 4570-4578.

Barry, S. P., S. M. Davidson and P. A. Townsend (2008). "Molecular regulation of cardiac hypertrophy." The International Journal of Biochemistry & Cell Biology **40**(10): 2023-2039.

Birol, I. (2009). "De novo transcriptome assembly with ABySS." Bioinformatics **25**: 2872-2877.

Chang, S., T. A. McKinsey, C. L. Zhang, J. A. Richardson, J. A. Hill and E. N. Olson (2004). "Histone Deacetylases 5 and 9 Govern Responsiveness of the Heart to a Subset of Stress Signals and Play Redundant Roles in Heart Development." Molecular and Cellular Biology **24**(19): 8467-8476.

Chen Gao, Y. W. (2012). "Global impact of RNA splicing on transcriptome remodeling in the heart." Journal of Zhejiang University Science B **13**(8): 603-608.

Deepak Srivastava, E. N. O. (2000). "A genetic blueprint for cardiac development." Nature **407**: 221-226.

Durocher, D., F. Charron, R. Warren, R. J. Schwartz and M. Nemer (1997). "The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors." EMBO J. **16**: 5687-5696.

Edwards, A. S. and A. C. Newton (1997). "Phosphorylation at Conserved Carboxyl-terminal Hydrophobic Motif Regulates the Catalytic and Regulatory Domains of Protein Kinase C." Journal of Biological Chemistry **272**(29): 18382-18390.

Eric Chruchill, G. B., Alice Vallentin, Tomoyoshi Koyanagi, and Daria Mochly-Rosen (2008). "PKC Isozymes in Chronic Cardiac Disease: Possible Therapeutic Targets." Annual Reviews Pharmacology and Toxicology(48): 569-599.

Gajewski, K., Y. Kim, Y. M. Lee, E. N. Olson and R. A. Schulz (1998). "D-Mef2: a target for tinman activation during Drosophila heart development." EMBO J. **16**: 515-522.

Gao, G. and S. C. Dudley Jr (2013). "RBM25/LUC7L3 Function in Cardiac Sodium Channel Splicing Regulation of Human Heart Failure." Trends in Cardiovascular Medicine **23**(1): 5-8.

Greulich, F., C. Rudat and A. Kispert (2011). "Mechanisms of T-box gene function in the developing heart." Cardiovascular Research **91**(2): 212-222.

Guo, W., S. Schafer, M. L. Greaser, M. H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A. M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. A. Hacker, K. W. Saupe, G. W. Dec, P. T. Ellinor, C. A. MacRae, B. Spallek, R. Fischer, A. Perrot, C. Ozcelik, K.

Saar, N. Hubner and M. Gotthardt (2012). "RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing." Nat Med **18**(5): 766-773.

Gusterson, R. J., E. Jazrawi, I. M. Adcock and D. S. Latchman (2003). "The Transcriptional Co-activators CREB-binding Protein (CBP) and p300 Play a Critical Role in Cardiac Hypertrophy That Is Dependent on Their Histone Acetyltransferase Activity." Journal of Biological Chemistry **278**(9): 6838-6847.

Hambleton, M., H. Hahn, S. T. Pleger, M. C. Kuhn, R. Klevitsky, A. N. Carr, T. F. Kimball, T. E. Hewett, G. W. Dorn, W. J. Koch and J. D. Molkentin (2006). "Pharmacological- and Gene Therapy-Based Inhibition of Protein Kinase Cα/β Enhances Cardiac Contractility and Attenuates Heart Failure." Circulation **114**(6): 574-582.

Harvey, R. P. (1996). Dev. Biol. **178**: 203-216.

Henderson, D. J. and A. J. Copp (1998). "Versican Expression Is Associated With Chamber Specification, Septation, and Valvulogenesis in the Developing Mouse Heart." Circulation Research **83**(5): 523-532.

Hong Ki Song, S.-E. H., Taeyong Kim, Do Han Kim (2012). "Deep RNA Sequencing Reveals Novel Cardiac Transcriptomic Signatures for Physiological and Pathological Hypertrophy." PLOS ONE.

Huang, Z.-P., J. Chen, H. Y. Seok, Z. Zhang, M. Kataoka, X. Hu and D.-Z. Wang (2013). "MicroRNA-22 Regulates Cardiac Hypertrophy and Remodeling in Response to Stress." Circulation Research **112**(9): 1234-1243.

Jingjing, L., Y. Xue, N. Agarwal and R. S. Roque (1999). "Human Müller cells express VEGF183, a novel spliced variant of vascular endothelial growth factor." Investigative Ophthalmology & Visual Science **40**(3): 752-759.

Joanna Kobylinska, W. D., Monika Cendrowska-Pinkosz, Anna Dworzanska, Teresa Hermanowics-Dryka, Joanna Kiszka, Elzbieta Staroslawska, Franciszek Burdan (2013). "Morphological and Moelcular Bases of cardiac development." Postepy Hig Med Dosw.

Jr, G. G. a. S. C. D. (2013). "SCN5A splicing variants and the possibility of predicting heart failure-associated arrhythmia." Expert Reveiw of Cardiovascular Therapy **11**(2): 117-119.

Kilic, A., A. Velic, L. J. De Windt, L. Fabritz, M. Voss, D. Mitko, M. Zwiener, H. A. Baba, M. van Eickels, E. Schlatter and M. Kuhn (2005). "Enhanced Activity of the Myocardial Na+/H+ Exchanger NHE-1 Contributes to Cardiac Remodeling in Atrial Natriuretic Peptide Receptor–Deficient Mice." Circulation **112**(15): 2307-2317.

Kuwahara, K., T. Nishikimi and K. Nakao (2012). "Transcriptional Regulation of the Fetal Cardiac Gene Program." Journal of Pharmacological Sciences **119**(3): 198-203.

Ladage, D., L. Tilemann, K. Ishikawa, R. N. Correll, Y. Kawase, S. R. Houser, J. D. Molkentin and R. J. Hajjar (2011). "Inhibition of PKCα/β With Ruboxistaurin Antagonizes Heart Failure in Pigs After Myocardial Infarction Injury." Circulation Research **109**(12): 1396-1400.

Lara-Pezzi, E., J. Gómez-Salinero, A. Gatto and P. García-Pavía (2013). "The Alternative Heart: Impact of Alternative Splicing in Heart Disease." Journal of Cardiovascular Translational Research **6**(6): 945-955.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research **109**(12): 1332-1341.

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

Liu, Q., X. Chen, S. M. MacDonnell, E. G. Kranias, J. N. Lorenz, M. Leitges, S. R. Houser and J. D. Molkentin (2009). "Protein Kinase Cα, but Not PKCβ or PKCγ, Regulates Contractility and Heart Failure Susceptibility: Implications for Ruboxistaurin as a Novel Therapeutic Approach." Circulation Research **105**(2): 194-200.

Lu, J., T. A. McKinsey, R. L. Nicol and E. N. Olson (2000). "Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases." Proceedings of the National Academy of Sciences **97**(8): 4070-4075.

Manetti, M., S. Guiducci, L. Ibba-Manneschi and M. Matucci-Cerinic (2011). "Impaired Angiogenesis in Systemic Sclerosis: The Emerging Role of the

Antiangiogenic VEGF165b Splice Variant." Trends in Cardiovascular Medicine **21**(7): 204-210.

Martin, J. A. and Z. Wang (2011). "Next-generation transcriptome assembly." Nat Rev Genet **12**(10): 671-682.

Merdzhanova, G., S. Gout, M. Keramidas, V. Edmond, J. L. Coll, C. Brambilla, E. Brambilla, S. Gazzeri and B. Eymin (2010). "The transcription factor E2F1 and the SR protein SC35 control the ratio of pro-angiogenic versus antiangiogenic isoforms of vascular endothelial growth factor-A to inhibit neovascularization in vivo." Oncogene **29**(39): 5392-5403.

Mjaatvedt, C. H., H. Yamamura, A. A. Capehart, D. Turner and R. R. Markwald (1998). "TheCspg2Gene, Disrupted in thehdfMutant, Is Required for Right Cardiac Chamber and Endocardial Cushion Formation." Developmental Biology **202**(1): 56-66.

Molnár, A., A. Borbély, D. Czuriga, S. M. Ivetta, S. Szilágyi, Z. Hertelendi, E. T. Pásztor, Á. Balogh, Z. Galajda, T. Szerafin, K. Jaquet, Z. Papp, I. Édes and A. Tóth (2009). "Protein Kinase C Contributes to the Maintenance of Contractile Force in Human Ventricular Cardiomyocytes." Journal of Biological Chemistry **284**(2): 1031-1039.

Murphy, L. L., A. J. Moon-Grady, B. F. Cuneo, R. T. Wakai, S. Yu, J. D. Kunic, D. W. Benson and A. L. George Jr (2012). "Developmentally regulated SCN5A splice variant potentiates dysfunction of a novel mutation associated with severe fetal arrhythmia." Heart Rhythm **9**(4): 590-597.

Olson, E. (2002). "A genetic blueprint for growth and development of the heart." Harvey Lect **98**: 41-64.

Orenes-Piñero, E., S. Montoro-García, J. V. Patel, M. Valdés, F. Marín and G. Y. H. Lip (2013). "Role of microRNAs in cardiac remodelling: New insights and future perspectives." International Journal of Cardiology **167**(5): 1651-1659.

Rana, M. S., V. M. Christoffels and A. F. M. Moorman (2013). "A molecular and genetic outline of cardiac morphogenesis." Acta Physiologica **207**(4): 588-615.

Sharma GD, N. H., Antonov AS, Gerrity RG, von Geldern T, Pandey KN (2002). "Expression of atrial natriuretic peptide receptor-A antagonizes the mitogen-activated protein kinases (Erk2 and P38MAPK) in cultured human vascular smooth muscle cells." Mol Cell Biochem **233**(2-3): 165-173.

Srivastava, D. (1997). "Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND." Nature Genet. **16**: 154-160.

Srivastava, D., P. Cserjesi and E. N. Olson (1995). "New subclass of bHLH proteins required for cardiac morphogenesis." Science **270**: 1995-1999.

Steinberg, S. F. (2008). "Structural Basis of Protein Kinase C Isoform Function." Physiological Reviews **88**(4): 1341-1378.

Taegtmeyer, H., S. Sen and D. Vela (2010). "Return to the fetal gene program." Annals of the New York Academy of Sciences **1188**(1): 191-198.

Tischer, E., D. Gospodarowicz, R. Mitchell, M. Silva, J. Schilling, K. Lau, T. Crisp, J. C. Fiddes and J. A. Abraham (1989). "Vascular endothelial growth

factor: A new member of the platelet-derived growth factor gene family." Biochemical and Biophysical Research Communications **165**(3): 1198-1206.

Turbendian, H. K., M. Gordillo, S.-Y. Tsai, J. Lu, G. Kang, T.-C. Liu, A. Tang, S. Liu, G. I. Fishman and T. Evans (2013). "GATA factors efficiently direct cardiac fate from embryonic stem cells." Development **140**(8): 1639-1644.

Ucar, A., S. K. Gupta, J. Fiedler, E. Erikci, M. Kardasinski, S. Batkai, S. Dangwal, R. Kumarswamy, C. Bang, A. Holzmann, J. Remke, M. Caprio, C. Jentzsch, S. Engelhardt, S. Geisendorf, C. Glas, T. G. Hofmann, M. Nessling, K. Richter, M. Schiffer, L. Carrier, L. C. Napp, J. Bauersachs, K. Chowdhury and T. Thum (2012). "The miRNA-212/132 family regulates both cardiac hypertrophy and cardiomyocyte autophagy." Nat Commun **3**: 1078.

Wahbi, K., V. Algalarrondo, H. M. Bécane, V. Fressart, C. Beldjord, K. Azibi, A. Lazarus, N. Berber, H. Radvanyi-Hoffman, T. Stojkovic, A. Béhin, P. Laforêt, B. Eymard, S. Hatem and D. Duboc "Brugada syndrome and abnormal splicing of SCN5A in myotonic dystrophy type 1." Archives of Cardiovascular Diseases(0).

Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nat Rev Genet **10**(1): 57-63.

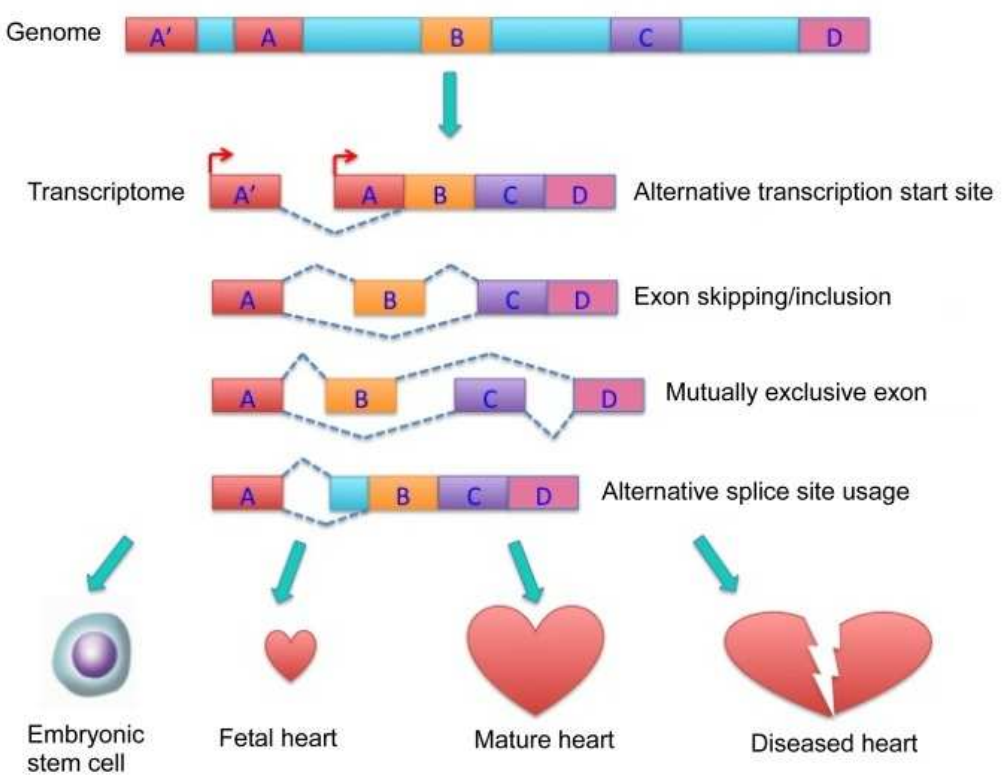
Wei Guo, S. J. B., Karla Esbona and Marion L. Greaser (2010). "Titin Diversity-Alternative Splicing Gone Wild." J Biomed Biotechnol.

Yanazume, T., K. Hasegawa, T. Morimoto, T. Kawamura, H. Wada, A. Matsumori, Y. Kawase, M. Hirai and T. Kita (2003). "Cardiac p300 Is Involved in Myocyte Growth with Decompensated Heart Failure." Molecular and Cellular Biology **23**(10): 3593-3606.

Zhang, C. L., T. A. McKinsey, S. Chang, C. L. Antos, J. A. Hill and E. N. Olson (2002). "Class II Histone Deacetylases Act as Signal-Responsive Repressors of Cardiac Hypertrophy." Cell **110**(4): 479-488.

Zhang, S.-S. and R. M. Shaw (2013). "Multilayered regulation of cardiac ion channels." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1833**(4): 876-885.

Figure 1. Graphic Abstract: Alternative Splicing and Cardiac Transcriptome Complexity



Chapter Two

Global Transcriptome Analysis in Pressure-Overload Induced Mouse Failing Heart

Introduction:

The development of RNA-Sequencing technology has allowed us to look at gene expression at single base resolution, which provided us novel insights into mammalian transcriptome. Previous study has identified extensive alternative splicing events and novel transcripts clusters in mammalian transcriptome. Although it has been suggested that cardiac genes also have alternative splicing events, including CamKinase, Tnnt and Titin, the scale of alternative splicing events in heart under normal and disease condition remains unexplored. In our RNA-Seq effort, we used transaortic constriction (TAC) to induce either early stage hypertrophy or late stage heart failure and compared with sham operated control hearts. We have developed two different methods including guided transcriptome reconstruction and de novo reconstruction to reconstruct transcript isoforms. In summary, we have identified 1435 genes showed differential expression between failing and normal hearts; we have detected a total of 7811 isoforms expressed in at least one sample, among these, a total of 720 isoforms in 475 genes were identified to be differentially expressed in different samples. We have also identified a total of 1873 novel exons corresponding to different types of alternative splicing events. Lastly, 1884 novel transcript clusters were identified in our sequencing data that do not overlap with any previous Ensemble genes, which have potential functional impact on cardiac disease based on their differential expression pattern between normal and failing hearts. Together, this deep RNA-Sequencing study has provided us novel insights into cardiac

transcriptome complexity during normal and disease conditions, which also set foundation for our future analysis on single splicing regulator and splicing event.

Circulation Research
JOURNAL OF THE AMERICAN HEART ASSOCIATION

American Heart Association®

Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts

Jae-Hyung Lee, Chen Gao, Guangdun Peng, Christopher Greer, Shuxun Ren, Yibin Wang and Xinshu Xiao

Circ Res. 2011;109:1332-1341; originally published online October 27, 2011;
doi: 10.1161/CIRCRESAHA.111.249433
Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7330. Online ISSN: 1524-4571

Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts

Jae-Hyung Lee, Chen Gao, Guangdun Peng, Christopher Greer, Shuxun Ren, Yibin Wang, Xinshu Xiao

Rationale: **Accurate and comprehensive de novo transcriptome profiling in heart is a central issue to better understand cardiac physiology and diseases. Although significant progress has been made in genome-wide profiling for quantitative changes in cardiac gene expression, current knowledge offers limited insights to the total complexity in cardiac transcriptome at individual exon level.**

Objective: **To develop more robust bioinformatic approaches to analyze high-throughput RNA sequencing (RNA-Seq) data, with the focus on the investigation of transcriptome complexity at individual exon and transcript levels.**

Methods and Results: **In addition to overall gene expression analysis, the methods developed in this study were used to analyze RNA-Seq data with respect to individual transcript isoforms, novel spliced exons, novel alternative terminal exons, novel transcript clusters (ie, novel genes), and long noncoding RNA genes. We applied these approaches to RNA-Seq data obtained from mouse hearts after pressure-overload–induced by transaortic constriction. Based on experimental validations, analyses of the features of the identified exons/transcripts, and expression analyses including previously published RNA-Seq data, we demonstrate that the methods are highly effective in detecting and quantifying individual exons and transcripts. Novel insights inferred from the examined aspects of the cardiac transcriptome open ways to further experimental investigations.**

Conclusions: **Our work provided a comprehensive set of methods to analyze mouse cardiac transcriptome complexity at individual exon and transcript levels. Applications of the methods may infer important new insights to gene regulation in normal and disease hearts in terms of exon utilization and potential involvement of novel components of cardiac transcriptome. (*Circ Res.* 2011;109:1332-1341.)**

Key Words: RNA-Seq ■ transcriptome profiling ■ hypertrophy ■ heart failure

Regulation of gene expression has a critical role in normal cardiac function and pathogenesis of heart failure. A global change in cardiac transcriptome from normal to one with characteristics of "fetal-like" profile is a major part of the pathological remodeling in failing hearts.[1–3] Although much insight has been learnt from transcriptome profiling studies using microarray technologies, limitations in coverage and sensitivity still leave a significant part of the cardiac transcriptome landscape unexplored, especially concerning expression and variation at single exon resolution. Recent advances in high-throughput sequencing technologies are enabling a new way to study transcriptomes: massively parallel sequencing of short reads derived from mRNAs (RNA-Seq).[4,5] Compared with microarray technologies, RNA-Seq was shown to enable more accurate quantification of gene expression levels.[6,7] More importantly, RNA-Seq does not require a priori annotation of gene and transcript structures. It allows not only in-depth studies of expression changes in known genes and alternative isoforms but also unbiased characterization of novel exons and novel transcript clusters. It also enables investigation of long noncoding RNA (lncRNA) genes, which are not usually targeted by alternative transcriptome profiling methods, such as microarrays. Thus, RNA-Seq opens the way to de novo transcriptome reconstruction and discovery of novel transcripts of any mammalian cell. Indeed, recent reports using RNA-Seq to profile transcriptome in mouse heart have revealed interesting new insights in cardiac transcriptional and signaling networks in genetic models of heart failure.[7–10]

In the present study, we developed bioinformatic methods to identify transcript structures and analyze transcriptome complexities with a particular emphasis on quantification of RNA splicing variants at single exon resolution using RNA-Seq data of normal and failing murine hearts. The methods

Original received May 27, 2011; revision received October 18, 2011; accepted October 20, 2011. In September 2011, the average time from submission to first decision for all original research papers submitted to *Circulation Research* was 16 days.

From the Department of Integrative Biology and Physiology (J.H.L., G.P., C. Greer, X.X.), the Molecular Biology Institute (C. Gao, Y.W., X.X.), the Departments of Anesthesiology (S.R., Y.W.), Physiology (S.R., Y.W.), and Medicine (S.R., Y.W.), and the Cardiovascular Research Laboratories (C. Gao, S.R., Y.W.), David Geffen School of Medicine, University of California, Los Angeles.

Data availability: RNA-Seq data are available at the Gene Expression Omnibus with ID GSE29446.

Correspondence to Xinshu Xiao, PhD, 611 Charles E. Young Drive, Boyer Hall, Room 660, Molecular Biology Institute, UCLA, Los Angeles, CA, 90095. E-mail gxxiao@ucla.edu, Phone: 310-206-6522, Fax: 310-206-9184

Circulation Research **is available at http://circres.ahajournals.org** **DOI: 10.1161/CIRCRESAHA.111.249433**

take full advantage of the strength of RNA-Seq. We show that they allowed in-depth profiling and quantification of alternative mRNA structures, novel exons, novel transcript clusters (NTCs) and long noncoding RNA genes. The results open ways to direct experimental investigation of these novel transcriptome features and highlight the power of RNA-Seq to provide a comprehensive bioinformatic delineation of disease-specific transcriptomes.

Non-standard Abbreviations and Acronyms	
AS	alternative splicing
ATSS	alternative transcription start site
HF	heart failure
HY	hypertrophy
lncRNA	long noncoding RNA
NTC	novel transcript cluster
RNA-Seq	RNA sequencing
RPKM	reads per kilobase of exon per million mapped reads
sham-HY	corresponding sham control for hypertrophy
sham-HF	corresponding sham control for heart failure

Methods

RNA-Seq Data Generation and Mapping

An expanded Methods section is available in the Data Supplement at http://circres.ahajournals.org.

Left ventricular tissues were collected from male C57BL/6 mice after 1 week (hypertrophy stage, HY) and 8 weeks after transaortic constriction (TAC) procedure (heart failure stage, HF) and their corresponding sham controls (sham-HY, sham-HF) (Online Supplement and Online Table I). To conduct RNA-Seq analysis, total RNAs from 6 TAC and sham-operated mice at the HY stage and 4 TAC and corresponding sham mice at the HF stage were obtained. Paired-end RNA-Seq reads (2×72nt or 2×76nt long) were mapped to the mouse Ensembl transcript sequences (release 56), using Bowtie[11] and BLAT.[12] Mapping was first carried out for individual reads without considering the read-pairing information. Next, all read pairs were inspected for correct pairing by considering whether they map to the same chromosome, potentially in the same gene and with correct orientation relative to each other (Online Supplement). The pair of reads was considered as uniquely mapped if and only if 1 unique pair of mapped locations was identified.

Analysis of Gene and Transcript Isoform Expression

Levels of gene and exon expression were quantified using the RPKM measure[13] and a minimum RPKM value of 3 (≈1 copy per cell) is required for expressed genes/isoforms (see Results for justification of this cutoff). Gene expression differences were evaluated using Fisher's exact test after normalizing by the total number of mapped reads in each lane using the upper-quartile normalization method.[14] The resulted probability values were corrected through the Benjamini and Hochberg method. Differentially expressed genes were defined as those with changes of at least 1.5-fold between a pair of samples at a false discovery rate (FDR) of 5% for genes expressed at ≥3 RPKM in ≥1 sample. The Cufflinks software (v 0.9.2)[15] was used to estimate expression levels of individual isoforms of an Ensembl gene, which allowed identification of isoform-specific expression changes due to alternative transcription start site (ATSS) or alternative splicing (AS). To further assess overall isoform expression dissimilarity in 2 samples (A and B), a dissimilarity score was defined based on the Morisita-Horn similarity index as follows:

$$dissimilary\ score = 1 - \frac{2\sum_i P_i(A)\,P_i(B)}{\sum_i [P_i^2(A) + P_i^2(B)]}$$

where $P_i(A)$ and $P_i(B)$ represent the expression of isoform i normalized by overall gene expression in the sample A or B. We only considered genes expressed at ≥3 RPKM in this analysis.

Transcriptome Reconstruction

We developed 2 different methods for the reconstruction of transcript isoforms: (1) guided transcriptome reconstruction: a method to reconstruct isoforms and discover novel exons within known genes; and (2) de novo reconstruction: a method to reconstruct completely new isoforms independent of known gene annotations. Details of the 2 approaches are presented in the Online Supplement. Briefly, the following steps are common to both approaches: (1) define expressed sequence fragments (seq-frags); (2) identify connections between seq-frags based on reads mapped to spliced junctions; (3) generate a directed graph using seq-frags and their connections for each gene or each chromosome; and (4) construct the isoforms by finding all possible paths in the graph. In guided transcriptome reconstruction, novel seq-frags can represent novel exons or extended regions of known exons. Identification of exon boundaries (novel or known) depends on the presence of reads mapped to spliced junctions. Thus, to reduce possible false-positive isoforms, we required at least 2 junction reads as evidence of a splicing event. In de novo reconstruction, NTCs were identified in intergenic regions and clustered together on each chromosome. The boundaries of NTCs were decided by confirming the absence of spliced junctions or expressed seq-frags. Filters for minimum expression levels of seq-frags and canonical splicing signals were applied. The coding potential of NTCs was evaluated using the Coding Potential Calculator software.[16]

Statistical and Computational Methods

For gene expression analysis, the statistical significance was assessed by Fisher's exact test as described above. Pearson correlation coefficients for gene expression validation were calculated in R. For GO analysis, empirical probability values were estimated based on 10 000 randomized simulations (Online Supplement), and the Bonferroni cutoff was used to determine significant probability values. All other computational procedures including transcript isoform reconstructions were carried out using in-house programs written in Python, Perl, and R.

Results

Mapping of RNA-Seq Reads

We obtained a total of 168 million pairs of reads using the standard paired-end RNA-Seq protocol on the Illumina GA II sequencer. Table 1 shows the number of reads in each TAC/sham group and the mapping results. Our mapping procedure (Figure 1) ensures that reads generated by both known genes and novel transcribed regions were identifiable. In addition, it enabled detection of novel and known spliced junctions that connect 2 or more exons intervened by long introns. The usage of paired-end sequencing brings the advantage of improved mapping performance. We estimated that 4% of all the original reads were mapped nonuniquely as singletons but uniquely as pairs. Ambiguous mapping results can be removed by examining the pairing of reads. For example, 12% of all reads were categorized as unmapped in the paired-end mode but mapped uniquely as singletons (possible mapping errors in the single-end mode). Only pairs of reads that mapped uniquely to the transcriptome and/or the genome were retained for further analyses (Figure 1). Among

Table 1. Number of RNA-Seq Reads Obtained for Each Type of Sample and Mapping Results

No. of Reads	HY	Sham-HY	HF	Sham-HF	Total
Total	83 718 570	83 009 180	100 602 620	68 442 422	335 772 792
Unmapped	15 325 416	12 007 844	36 768 368	23 961 060	88 062 688
Nonuniquely paired	10 483 514	11 559 470	6 689 996	5 398 104	34 131 084
Wrong pairing	6 405 098	6 194 896	6 470 068	4 359 302	23 429 364
Total uniquely paired (ie, final mapped)	51 504 542	53 246 970	50 674 188	34 723 956	191 049 656
Mapped within exons	37 882 946 (73%)	39 198 116 (73%)	32 479 491 (64%)	22 977 952 (66%)	132 538 505 (70%)
Mapped to exon-exon junctions	9 570 514 (19%)	9 977 970 (19%)	10 627 444 (21%)	7 442 858 (21%)	37 618 786 (20%)
Mapped to introns	1 619 022 (3%)	1 555 896 (3%)	3 058 379 (6%)	1 688 736 (5%)	7 922 033 (4%)
Mapped to intergenic regions	2 432 060 (5%)	2 514 988 (5%)	4 508 874 (9%)	2 614 410 (8%)	12 070 332 (6%)

HY indicates hypertrophy; HF, heart failure.

Percentages shown are relative to the number of final mapped reads.

"Unmapped" reads refer to those that were not mappable to the genome or transcriptome using the defined mismatch thresholds (Online Supplement). "Nonuniquely paired" means that the pair of reads mapped nonuniquely as a pair. "Wrong pairing" means that the pair of reads did not pass the filters for correct pairing (Online Supplement).

the 168 million pairs of reads obtained in our study, about 95 million (57%) were uniquely mapped in total, which covered 72% of the known exon-exon junctions, 82% of exon bodies and 77% of known genes.

Analysis of Gene Expression Levels Through RNA-Seq

RNA-Seq has been demonstrated to be an effective approach for gene expression profiling in mouse heart.[7–9] One advantage of this method is its ability to provide quantitative read-out of the mRNA expression levels in 1 sample, in contrast to microarrays that just permit comparative analyses without absolute expression values. Consistent with the previous studies,[7–9] we also observed a wide dynamic range of expression values varying from about 1 copy per cell (3 RPKM) for the *Ankrd12* gene to 8,048 copies per cell for the *mt-Co1* gene in the sham hearts. In this work, we use 3 RPKM as the minimum cutoff to filter for expressed genes[13] considering its biological relevance and the fact that some genes (eg, *Kcnd2, Kcnd3*) with heart-related function are expressed at about 3 RPKM in shams. Other RNA-Seq studies also showed that low abundant transcripts expressed at about 1 copy per cell include transcriptional factors and other functionally important genes for cardiac regulation.[7–9] In addition, in a PubMed search of published abstracts, we found that genes with ≥3 RPKM expression levels are about twice as often associated with the keywords "heart" and "cardiac" than those expressed at <3 RPKM. Altogether, 9833 genes (29% of all Ensembl genes) are expressed at ≥3 RPKM in at least 1 sample in our study.

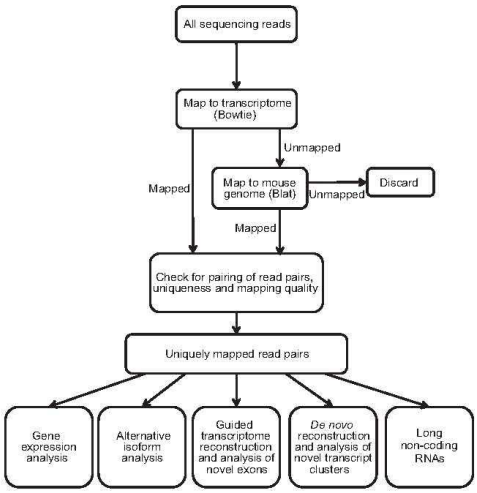


Figure 1. Overview of the methods and procedures to analyze RNA-Seq data.

Another advantage of the RNA-Seq method over microarrays is the improved quantification of differential gene expression between samples.[7] We validated expression changes of 42 genes using real-time PCR (Online Figure I and Online Table II). This validation demonstrated that the results from RNA-Seq and real-time PCR are highly concordant (r=0.90, Pearson correlation). To establish biologically meaningful criteria to determine significant differential gene expression, we examined the levels of genes known to be altered in failing hearts, including *Myh7*, *Egr1*, *Nppb*, *Pln*, and *Actb*. We confirmed that all the above genes had expression changes of at least 1.5-fold between the HF and sham-HF samples in the real-time PCR or RNA-Seq. Thus, we used the following criteria to identify differentially expressed genes: (1) gene expression level ≥3 RPKM in either sham or HY/HF or both; (2) change in expression level ≥1.5 fold; and (3) Fisher's exact test (see Methods) FDR <5%. Altogether, 97 and 1435 genes passed the above filters between sham-HY and HY, sham-HF and HF, respectively (Online Table III). The number of genes and their magnitudes of changes are larger at the HF stage, consistent with the expected higher degrees of cardiac remodeling in HF contrib-

uted by vascular remodeling, inflammatory response and fibrosis in the myocardium. In addition to the well-known HF or HY-related genes, we found genes with significant alteration in expression but that have never been implicated in heart failure or cardiac hypertrophy, such as *Prc1, E2f1, Birc5, Iqgap3, Cdc20,* and *Cdca8*. Our results suggest that RNA-Seq may provide novel insights in overall gene expression as demonstrated previously.[7–9]

Analysis of Alternative Transcript Isoforms Through RNA-Seq

One of the main voids in our current knowledge of cardiac transcriptome is the genome-wide profile of mRNA splicing variants, a very challenging task not readily accomplishable by most microarray platforms. For this purpose, we used the package Cufflinks,[15] which was shown to effectively capture isoform-specific expression and alteration in RNA-Seq data. Although Cufflinks has a number of modules for different purposes, we focused on its usage to infer expression levels and differential expression of individual transcript isoforms. As inputs to Cufflinks, we used our read mapping results described above and the set of Ensembl-defined genes and their spliced variants. The novel exons and NTCs identified in our study were not included because the nature of the short sequencing reads limits the accuracy in predicting complete structures of spliced variants that are needed to estimate their expression levels.

We first examined the absolute isoform expression estimated by Cufflinks. For genes with multiple transcript isoforms, we detected a total of 7811 isoforms with expression level ≥3 RPKM in at least 1 sample. The most abundant isoform was from the gene *Myl2* (Myosin regulatory light chain 2) in the sham-HF sample. The isoforms of other heart-related transcripts such as *Actc1*, *Atp2a2*, *Myh6, Tnnt2,* and *Tpm1* were also highly expressed. We observed spliced variants for numerous genes (eg, *Atp2a2*, *Cacna1c*, *Slc6a8,* and *Ank2*) known to undergo alternative splicing in cardiac tissues.[7] For the gene *Camk2*, we confirmed that its neuronal-specific isoform is much less expressed than the heart-specific isoforms (2.96 versus 22.11 RPKM) in the sham-HF sample. These findings suggest that RNA-Seq can readily detect isoforms of a gene due to alternative RNA splicing.

We next analyzed the differential expression patterns of individual transcript isoforms. Cufflinks analysis identified 1087 genes (mostly protein-coding genes) with significant isoform-specific expression changes (q-value <0.05) due to ATSS or AS or both (Figure 2A and Online Table IV). As examples, Figure 2B shows 2 genes (with ATSS and AS, respectively), their reads distributions, and RT-PCR validation results. If the same criteria were applied as for determining differential gene expression (q-value <0.05, fold change ≥1.5, and expression level ≥3 RPKM), a total of 720 isoforms in 475 genes were identified as differentially expressed. Overall, genes with ATSS and AS are both enriched in biological processes related to muscle function and ATP synthesis (Online Table V).

Similar to our findings in overall gene expression changes, more genes were found to have altered isoform expression in the HF stage than the HY stage. To further compare the 2 stages, we calculated a dissimilarity score that quantifies the overall isoform difference of a gene between a pair of samples (Methods). This measure is independent of gene expression levels. Figure 2C shows that most genes have similar scores between the HY and HF stages (data distributed close to the diagonal line). However, a significant number of genes (n>250) have dissimilarity scores differing by more than 0.2, suggesting a significant change in isoform usage at different stages of heart failure. Interestingly, among these genes, many (eg, *Garnl1, Sipa1, Rgs12, Rin2,* and *Rabgap1*) are known to be involved in processes well-studied in heart failure. In addition, genes involved in chromatin and histone modifications (such as *Hdac7, Ezh1,* and *Aof2*) also demonstrated stage-specific isoform expression changes. Therefore, the quantitative gene isoform analysis suggests a global change in exon utilization due to alternative RNA splicing that can potentially affect functionally important genes in failing hearts.

Guided Transcriptome Reconstruction for Novel Transcripts of Known Genes

Another major advantage of the RNA-Seq approach is the capability to discover previously unknown transcript isoforms. We thus developed a guided transcriptome reconstruction method to enable identification of novel isoforms in known genes (Online Supplement). In this method, RNA-Seq reads that mapped inside or in the vicinity of Ensembl genes were examined. Reads that do not support the known Ensembl transcript structures (eg, those in the intronic regions) may suggest existence of novel transcripts. However, such reads may also arise from other sources such as incompletely processed transcripts, degradation intermediates of introns, or mapping errors. To reduce false-positives, we implemented 2 additional requirements to define novel transcripts. First, we applied stringent filters for expression levels of the novel fragments (details in Online Supplement). Second, because novel exons should be spliced to other exons, we required the existence of at least 2 spliced junction reads flanking each end of the novel exon. When the detected novel fragments were identified to be extensions of known internal exons or terminal exons, they would indicate new exon splicing pattern, or alternative transcriptional initiation or termination events.

For the 10 061 multi-exon genes (defined by Ensembl and/or our isoform reconstruction) with an expression level of ≥3 RPKM, 5112 (51%) were detected with novel isoforms (with novel exons or novel splicing patterns among known exons) as a result of alternative splicing, 1651 (16%) as a result of alternative initiation or termination, and 830 (8%) with both types of novel isoforms. Novel transcript structures were identified across a broad range of expression levels (Online Figure II), with more isoforms detected for higher expressed genes (most likely due to higher read coverage). Therefore, our findings suggest a significant deficiency in the current mouse transcriptome annotation (Ensembl v 56 used in this analysis) and deep RNA-Seq combined with guided transcriptome reconstruction can provide a much more comprehensive profile of the total complexity in transcript structures.

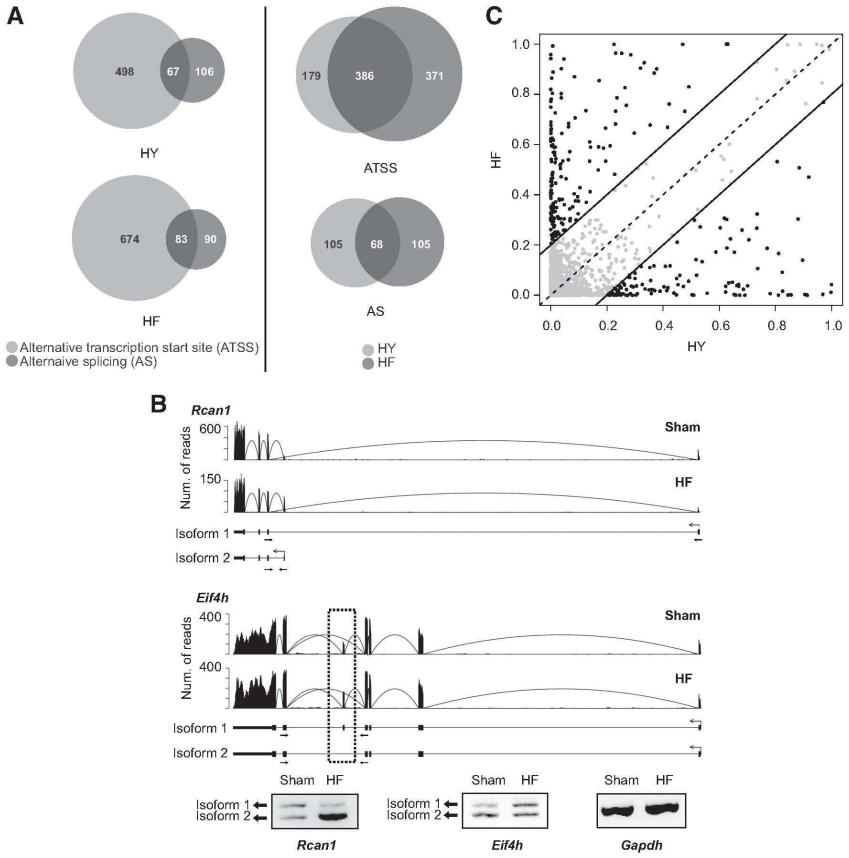


Figure 2. Expression and isoform changes of known genes in heart failure. A, Number of genes with expression changes of alternative isoforms due to alternative transcription start sites (ATSS) or alternative splicing (AS). **Left panel:** genes classified into HY and HF stages; **right panel:** genes classified into the ATSS and AS categories. **B**, Read distribution for 2 genes with differential isoform expression due to ATSS (*Rcan1*) or AS (*Eif4h*) and the corresponding RT-PCR validation results (with primer locations illustrated by small arrows). Arcs represent reads mapped to exon-exon junctions. Ensembl-annotated isoforms are illustrated below read distributions. In the *Eif4h* gene, the skipped exon is highlighted by a **dotted box**. **C**, Distribution of dissimilarity scores to quantify the overall isoform difference of genes in the 2 stages relative to their sham samples.

Evaluation of Novel Spliced Exons Identified by Guided Transcriptome Reconstruction

Although a large number of genes were identified with novel spliced variants, it is possible that many of them were resulted from random errors or noise in the process of splicing or transcription detected by the highly sensitive RNA-Seq method. Thus, we analyzed in detail the novel exons in the spliced variants to determine if our method leads to findings with potential biological significance. A total of 1873 novel exons were identified corresponding to different types of alternative splicing events (Table 2 and Online Table VI). Because the Ensembl v 56 database was used as a reference to define known genes and exons and updated databases now exist, we examined whether the above novel exons are annotated as known exons in the new Ensembl v 61 database (April 2011), or the UCSC, RefSeq databases. Indeed, 26% of the novel spliced exons identified from our original analysis are now "known" to one or more of the above databases (Table 2). This serves as a validation of our approach for identifying the novel exons and we refer to those exons that remain to be not annotated in the above databases (1384 in total) as "updated novel exons." We validated 29 of the randomly selected updated novel exons through RT-PCR and the expression of 97% of them was confirmed (Online Table VII).

To provide further evaluation on the identified "updated novel exons," we performed additional bioinformatic analyses on these novel exons that were alternatively skipped, the most common type of alternatively spliced exons. We first

Table 2. Novel Alternative Splicing Events Identified in Ensembl Genes and Novel Transcript Clusters

AS Events	SE*	RI†	A5E‡	A3E§	MXE‖
Original novel exons in ENSG	968	691	268	332	15
Percentage of total (1873)	52%	37%	14%	18%	<1%
Updated novel exons in ENSG	623	629	197	235	8
Percentage of total (1384)	45%	45%	14%	17%	<1%
Novel AS exons in NTCs	223	0	138	109	7
Percentage of total (421)	53%	0	33%	26%	1%

Exons are categorized according to the type of alternative splicing (AS) events. "Original" indicates novel exons identified relative to Ensembl v 56. "Updated" indicates novel exons identified relative to the most recent databases including Ensembl v 61, UCSC KnownGenes, and RefSeq genes. Note that 1 exon may be associated with multiple types of AS. Such exons are counted into all applicable types. Thus, the percent values for all categories may not sum to 100%. ENSG indicates Ensembl genes; NTC, novel transcript clusters.

*SE: Skipped exon.

†RI: Retained intron.

‡A5E: Alternative 5′ splice site exon.

§A3E: Alternative 3′ splice site exon.

‖MXE: Mutually exclusive exon.

examined whether the novel exons have features that resemble those of known skipped exons. The following features were considered: evolutionary conservation, expression level, exon length, and splice site strength. The conservation level and expression level of the novel exons, although lower than those of the known skipped exons, are significantly higher than intronic regions with matched GC content and length (Figure 3A and 3B). About 54% of the novel skipped exons have a length that is multiples of 3, significantly higher than expected (probability value=3.1e-07, χ^2 test). This observation implies the existence of strong selection on the alternative protein products derived from these exons. Approximately 96% of the novel skipped exons are flanked by GT-AG in the immediate intronic regions, representing consensus splice site sequences. Figure 3C shows the result of principal component analysis of the above features in the novel and the known exon populations, respectively, which suggests that the 2 groups of exons are largely similar. The similarities of the novel exons to the known alternative exons suggest that they are likely authentic exons with biological function.

To further evaluate their biological significance, we then analyzed the expression patterns of the novel exons in more detail. If a novel exon exists due to nonfunctional random splicing noise, its absolute expression level is most likely low and similar across different samples. However, we found about 72% of the novel exons had an expression level of ≥3 RPKM in at least 1 of the samples used in our study. In examining the expression difference of the novel exons in the HY/HF and sham controls, we observed that a substantial fraction (682 exons, 68% of all with ≥3 RPKM in absolute expression) had an expression difference of at least 1.5-fold (10 examples shown in Figure 3D, left panel). Furthermore, we computed the expression of the novel exons in other mouse tissues or cell types based on available RNA-Seq data.[13,17,18] Interestingly, many novel exons with relatively low abundance in the mouse heart had much higher expression levels in one or more of other tissues (Online Figure III; examples in Figure 3D). Thus, our result suggests that the novel exons identified by RNA-Seq even with low expression level in heart may be authentic exons with biological roles in other tissues.

Finally, we analyzed the impact of these novel exons on the predicted protein products. A large fraction (60%) of them will introduce premature termination codons or are expected to induce nonsense-mediated decay, suggesting that many novel exons may have a major impact on the final protein expression (Online Supplement). Furthermore, we found evidence of translation for 174 of the detected novel exons in public proteomic databases, suggesting that these exons may contribute to the overall complexity of the proteome (Online Supplement).

Alternative Initiation and Termination Events

On the basis of the reconstructed transcript structures, we identified novel alternative 5′ terminal exons that differ from the Ensembl annotations. Two different categories of such events were defined: (1) 5′ terminal exon overlapping the annotated first exon but with extended regions or alternative splice sites supported by junction reads and (2) 5′ terminal exon not overlapping any annotated exon and occurring upstream of the annotated 5′ start sites. To be conservative and avoid the complication of incomplete transcript reconstruction, we excluded 5′ terminal exons whose start sites are downstream of the annotated 5′ start sites. Similar analyses were carried out to identify alternative 3′ terminal exons. Altogether, we identified 1613 exons (in 1535 genes expressed at ≥3 RPKM) with novel alternative initiation or termination events that are not annotated in the most recent databases including UCSC, RefSeq and Ensembl (v 61) (Online Tables VIII and IX). Figure 4 shows 2 examples of such events. Among these exons, about 469 differ in expression by at least 1.5-fold between HY and sham or HF and sham. In addition, the corresponding genes significantly overlap the list of differentially expressed genes (162 genes in common, P=4.2e-13). Interestingly, 39% of the alternative 3′ terminal exons contain predicted target sites of known miRNAs expressed in mouse heart (Online Supplement), suggesting that the alternative terminal events may be functionally involved in gene regulation.

Methods to Identify and Evaluate Novel Transcript Clusters

In addition to the guided transcriptome reconstruction, we also conducted de novo transcript identification for those reads that match to genomic regions with no annotated genes (Methods). We focused on NTCs corresponding to genes with multiple exons only. We refer to an NTC as a cluster of all possible transcript isoforms. It is possible that an NTC contains false-positive isoforms because the nature of the short reads in RNA-Seq does not allow identification of complete transcript structures. In addition, complete profile of all isoforms may not be identifiable due to low read coverage. Nevertheless, each NTC suggests the possible existence of a novel gene with multiple transcript isoforms.

Using these criteria, 1884 NTCs were identified that do not overlap any Ensembl genes (v 56), all of which were multi-exon transcripts (Online Table X). Among the 5869 exons within these clusters, 8% had spliced junction reads

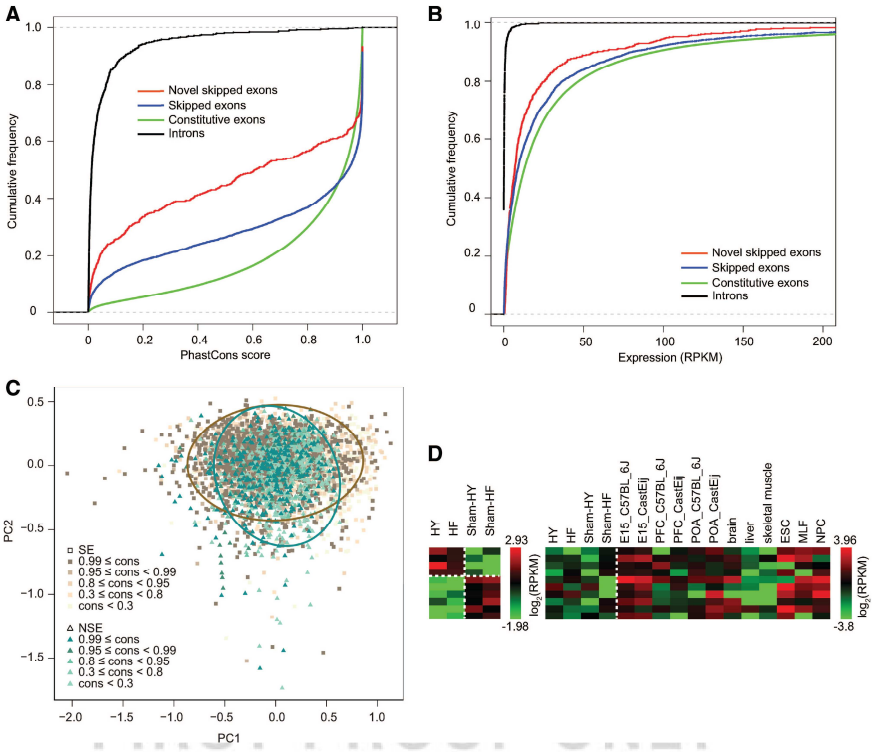


Figure 3. Novel exons in known genes identified through RNA-Seq. A, Cumulative distribution functions of conservation levels of novel skipped exons, introns, known skipped exons, and constitutive exons. **B**, Cumulative distribution functions of expression levels of the same exons/introns in **A**. **C**, Principal component analysis (PCA) and clustering of novel (NSE) and known skipped exons (SE), based on their expression level, exon length, and splice site strength. Conservation level (cons) of the exonic regions is represented in varying shades of colors. The representative ellipses for novel and known skipped exons were obtained by fitting the principal components with Gaussian distributions (Online Supplement). **D**, Heatmaps of expression levels (log_2RPKM; reads per kb of exon per million mapped reads) of novel skipped exons. **Left panel:** 10 example novel exons with differential expression in HY or HF compared with sham controls. **Right panel:** 10 example novel exons lowly expressed in our data, but highly expressed in one or more of the other published mouse RNA-Seq data sets. C57BL_6J and Cast/Eij are mouse strains used in.[17] E15 indicates embryonic day 15 (E15) whole brain; PFC, adult male and female medial prefrontal cortex; POA, adult male and female preoptic area.[17] Brain, liver, and skeletal muscle data from Mortazavi et al.[13] ESC indicates embryonic stem cells; MLF, lung fibroblasts; and NPC, neural progenitor cells.[18]

suggesting alternative splicing, with the most prevalent type being alternatively skipped exons (Table 2). Of all NTCs, 863 (46%) are now annotated as known genes in the most recent databases of UCSC, Refseq and Ensembl (v 61), supporting the validity of our method in identifying novel genes. We validated the expression patterns of 7 NTCs (3 newly annotated and 4 remain novel) using real-time PCR in the same mouse samples as used for RNA-Seq (Online Figure I and Online Table II). The results confirmed the expression of all 7 NTCs and also showed that RNA-Seq and real-time PCR gave highly consistent measures in gene expression changes (r=0.87, Pearson correlation).

Next, we focused on the 1021 NTCs that are still not annotated in the newest databases. To infer biological significance, we analyzed their expression patterns in our samples and the other mouse RNA-Seq data sets mentioned above.[13,17,18] Remarkably, many of these NTCs with low abundance in the heart had much higher expression levels in other tissues (examples shown in Figure 5A). A total of 199 NTCs passed the minimum expression level cutoff of 3 RPKM in at least 1 sample. When examined for differential expression using the same criteria as for known genes, 195 NTCs were differentially expressed between at least 1 pair of samples (34 between HF/HY and sham controls). In addition, we found that the NTCs

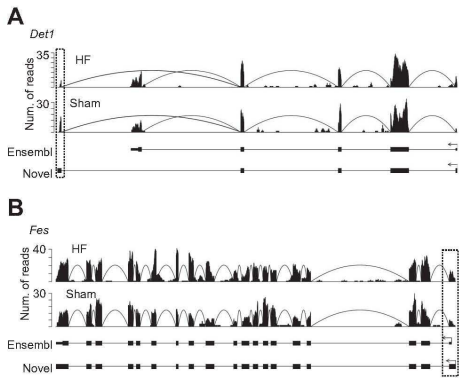


Figure 4. Examples of novel alternative terminal exons identified through RNA-Seq. Read distribution plots defined similarly as in Figure 2. Novel terminal exons are highlighted in **dotted boxes.** The Ensembl-annotated isoforms (Ensembl) and the novel isoforms (Novel) reconstructed from RNA-Seq are shown. **A**, Novel alternative 3′ terminal exon in the gene *Det1*. **B**, Novel alternative 5′ initiation exon in the gene *Fes*.

may be more tissue-specific than annotated genes evaluated by an entropy-based tissue specificity index (Online Supplement and Figure 5B). These results suggest that many NTCs may be actively regulated and may have functional ramifications in specific tissues.

We next classified the NTCs into coding and noncoding clusters. Among all 1021 NTCs, 105 clusters (containing 315 possible transcript isoforms) were found to have significant coding potential.[16] Interestingly, the coding clusters had significantly higher expression and conservation levels than the noncoding clusters (Online Figure IV), consistent with the existence of stronger selection pressure on protein-coding sequences as observed in known genes. Indeed, we found that 46% (48 of 105) of the putatively coding NTCs had sequence matches in public proteomic databases (Online Supplement). These results suggest that our de novo transcript identification method can effectively identify novel transcripts from RNA-Seq data sets, and the novel genes identified in heart contribute to the total complexity of cardiac transcriptome and proteome.

Inference of Potential Function of Novel Transcript Clusters

Functionally related genes involved in the same biological pathways or protein interaction networks are often regulated by similar transcription factors or other gene regulators. Thus, one approach to infer the potential function of novel genes is by determining whether their expression patterns correlate with those of known genes of certain function, based on coexpression analysis.[19] Note that such analyses only provide tentative functional indications of a gene. However, they may help to formulate hypotheses for further experimental studies. We applied this scheme to examine the potential functions of NTCs. Using the WGCNA method,[20] we constructed coexpression networks encompassing all known genes and 98 NTCs discovered in the heart samples (≥3 RPKM). We identified 52 network modules in total and focused on 20 that are enriched with differentially expressed genes between HY/HF and the corresponding shams. Genes in these modules are significantly associated with GO categories related to heart and muscle functions (Online Table XI). A total of 58 NTCs were included in the 20 significant modules. Among them, we analyzed the 10 most highly connected NTCs (ie, with highest connectivity) in each module that had at least 10 neighboring known genes. Fifteen NTCs from 6 modules were chosen in this way (example shown in Figure 6A). A complete list of GO categories related to each of the 15 NTCs is included in Online Table XII.

Long Noncoding RNA Genes

Among the novel multi-exon transcript clusters, a large fraction (81%) had low coding potential (Online Table X) and thus most likely belongs to the category of lncRNA genes. LncRNAs were recently recognized to have diverse functions in gene regulation and may contribute to disease etiology.[21] To assess the expression of lncRNAs in mouse heart failure, we combined our noncoding NTCs with the Ensembl and other mouse lncRNA data sets[18,22] to constitute a comprehensive repository of 5767 lncRNAs. A total of 703 lncRNAs were expressed at levels ≥3 RPKM in our RNA-Seq data of at least 1 sample. Interestingly, for lncRNAs expressed below 3 RPKM, 62% of them had an expression level of at least 3 RPKM in 1 or more of the publicly available mouse RNA-Seq data.[13,17,18] Figure 5B shows that lncRNAs have higher tissue specificity in their expression than protein-coding genes on average. Among all expressed lncRNAs (≥3 RPKM in ≥1 of our samples), 15 and 135 are differentially expressed (≥1.5-fold change, FDR <5%) between HY and sham-HY and between HF and sham-HF, respectively. Intriguingly, the well-known *H19* gene demonstrated significant upregulation in the HF stage compared with the corresponding sham controls (Figure 6B). This gene is highly expressed during embryogenesis and was shown to have tumor-suppressor activity.[23] However, its role in heart failure is not yet clear.

Discussion

Accurate and de novo transcriptome profiling is a central issue in disease research. We describe methods and applications of transcriptome analysis through RNA-Seq in normal and failing murine hearts. As clearly demonstrated by Matkovich et al,[7] RNA-Seq has both accuracy and sensitivity to detect transcripts with low abundance (such as those encoding transcription factors), so new gene expression networks associated with heart failure can be established (such as transcriptional networks). Our work highlights the power of RNA-Seq in de novo transcriptome profiling given its accuracy and the independence of a priori knowledge of transcripts to be analyzed. We used newly developed bioinformatic tools, namely a combination of guided transcriptome reconstruction and de novo reconstruction approaches, to identify novel exons and transcripts in normal and diseased mouse hearts. As demonstrated by our validation studies, the vast majority of the identified novel exons or transcripts are confirmed by RT-PCR for their expression in heart. We

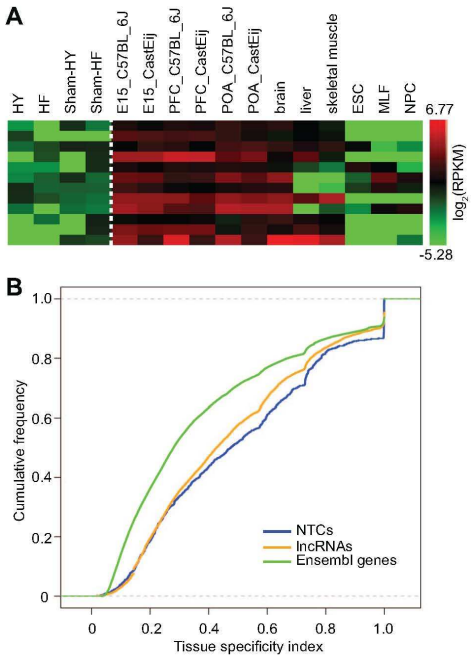


Figure 5. Analysis of novel transcript clusters (NTCs) and long noncoding RNA (lncRNA) genes identified through RNA-Seq. A, Heatmap of expression levels (log_2RPKM) of 12 example NTCs in our data and other published RNA-Seq data sets; see Figure 3 for details of the published data sets. **B**, Cumulative distribution functions of the tissue-specificity index of NTCs, all novel and known lncRNA genes and Ensembl protein-coding genes.

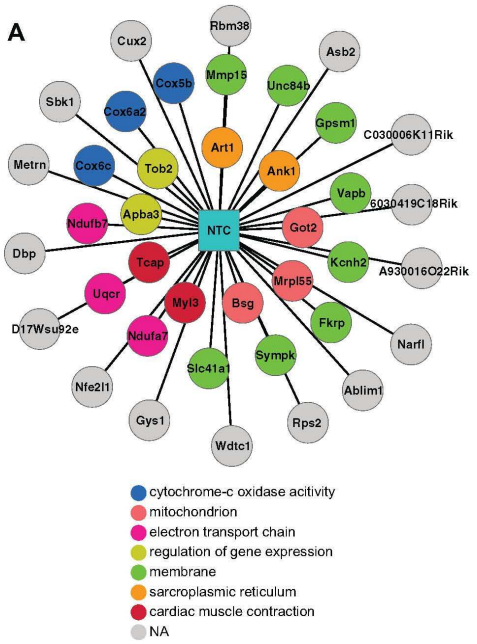


Figure 6. Novel transcript clusters (NTCs) and long noncoding RNA (lncRNA) genes potentially involved in disease mechanisms. A, An example NTC and its neighboring genes in the gene coexpression networks. Nongray nodes correspond to genes that are associated with GO categories enriched among all the neighboring genes of the NTC. Otherwise, the nodes are denoted in gray. **B**, Read distributions of a differentially expressed lncRNA gene (*H19*) in HF and sham RNA-Seq data.

showed that the sequence and conservation features of the novel exons are generally similar to those of the known alternatively spliced exons. In addition, both novel exons and NTCs were found to be expressed in a tissue-specific manner. These properties suggest the potential functional significance of the novel isoforms and novel transcripts. They also suggest that the RNA-Seq technology, combined with bioinformatic analysis, is a sensitive tool to comprehensively profile the transcriptome complexity at individual exon resolution.

It is somewhat unexpected to us that cardiac transcriptome encompasses such a large number of novel transcripts and exons with potential biological significance that have never been annotated despite extensive genome-wide profiling of cardiac transcriptome in the past decade. These findings open the way to further experimental investigations of their relevance in the pathogenesis of heart failure. Yet, our study does have limitations. Because the findings are mainly based on bioinformatic observations, their functional relevance would need to be further established at molecular and cellular levels experimentally. In addition, the poly-A selection procedure during RNA-Seq library preparation may introduce a positional bias in the coverage of the entire transcript (favoring the 3′ end) and not all expressed RNAs can be included in the library. In this study, we focused on pressure-overload–induced heart failure, and whether the findings are general to other types of heart failure models is unclear. Nevertheless, our analyses revealed previously uncharacterized complexity in the cardiac transcriptome as well as their dynamic changes during heart failure. This study highlights the need to use both RNA-Seq and bioinformatic tools to reevaluate cardiac transcriptome in other heart disease models as well as in human heart failure. Because the bioinformatic approaches developed in our study are generally applicable to any RNA-Seq data sets, we suggest that application of these new tools will vastly expand our current knowledge of transcriptome architecture and dynamics in general.

Sources of Funding

This work was supported in part by NIH grant R21DA027039, an Alfred P. Sloan Foundation Research Fellowship, and research grant

No. 5-FY10-486 from the March of Dimes Foundation to X.X., an American Heart Association Postdoctoral Fellowship to J.H.L., and NIH grants HL088640, HL070079, HL103205, and HL098954 to Y.W.

Disclosures

None.

References

1. Olson EN. Gene regulatory networks in the evolution and development of the heart. *Science*. 2006;313:1922–1927.
2. Olson EN, Backs J, McKinsey TA. Control of cardiac hypertrophy and heart failure by histone acetylation/deacetylation. *Novartis Found Symp*. 2006;274:3–12; discussion 13–19, 152–155, 272–276.
3. Hill JA, Olson EN. Cardiac plasticity. *N Engl J Med*. 2008;358:1370–1380.
4. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10:57–63.
5. Garber M, Grabherr MG, Guttman M, Trapnell C. Computational methods for transcriptome annotation and quantification using RNA-Seq. *Nat Methods*. 2011;8:469–477.
6. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-Seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*. 2008;18:1509–1517.
7. Matkovich SJ, Zhang Y, Van Booven DJ, Dorn GW II. Deep mRNA sequencing for in vivo functional analysis of cardiac transcriptional regulators: application to Galphaq. *Circ Res*. 2010;106:1459–1467.
8. Zhang Y, Matkovich SJ, Duan X, Gold JI, Koch WJ, Dorn GW II. Nuclear effects of g-protein receptor kinase 5 on histone deacetylase 5-regulated gene transcription in heart failure. *Circ Heart Fail*. 2011;4:659–668.
9. Zhang Y, Matkovich SJ, Duan X, Diwan A, Kang MY, Dorn GW II. Receptor-independent protein kinase C alpha (PKCalpha) signaling by calpain-generated free catalytic domains induces HDAC5 nuclear export and regulates cardiac transcription. *J Biol Chem*. 2011;286:26943–26951.
10. Xiang SY, Vanhoutte D, Del Re DP, Purcell NH, Ling H, Banerjee I, Bossuyt J, Lang RA, Zheng Y, Matkovich SJ, Miyamoto S, Molkentin JD, Dorn GW II, Brown JH. RhoA protects the mouse heart against ischemia/reperfusion injury. *J Clin Invest*. 2011;121:3269–3276.
11. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10:R25.
12. Kent WJ. BLAT–the BLAST-like alignment tool. *Genome Res*. 2002;12:656–664.
13. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5:621–628.
14. Bullard JH, Purdom E, Hansen KD, Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics*. 2010;11:94.
15. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28:511–515.
16. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007;35(Web Server issue):W345–W349.
17. Gregg C, Zhang J, Weissbourd B, Luo S, Schroth GP, Haig D, Dulac C. High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science*. 2010;329:643–648.
18. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, Koziol MJ, Gnirke A, Nusbaum C, Rinn JL, Lander ES, Regev A. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol*. 2010;28:503–510.
19. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. *Science*. 2003;302:249–255.
20. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
21. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev*. 2009;23:1494–1504.
22. Ponjavic J, Oliver PL, Lunter G, Ponting CP. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. *PLoS Genet*. 2009;5:e1000617.
23. Gabory A, Jammes H, Dandolo L. The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays*. 2010;32:473–480.

Novelty and Significance

What Is Known?

- Accurate and de novo transcriptome profiling is a central issue in studying the mechanisms of cardiovascular development and diseases.
- Understanding of the global cardiac transcriptome landscape is currently limited concerning expression and variation at single exon resolution.
- Whole-transcriptome sequencing (RNA-Seq) offers a new way to study transcriptomes.

What New Information Does This Article Contribute?

- We present bioinformatic methods to identify transcript structures and to analyze transcriptome complexities with a particular emphasis on quantification of RNA splicing variants at single exon resolution using RNA-Seq data of normal and failing murine hearts.
- We validate the effectiveness and accuracy of our bioinformatic approaches, based on experimental confirmation and cross-database analyses.
- We show that the bioinformatic analyses of RNA-Seq allow in-depth profiling and quantification of alternative mRNA structures, novel exons, novel transcript clusters and long noncoding RNA genes in mouse heart.

Transcriptome profiling offers detailed insight in understanding gene regulation in health and diseases. Previous technologies (eg, microarrays) for this purpose are limited in coverage and sensitivity. RNA-Seq using massively parallel next-generation sequencing platforms can potentially enable de novo and unbiased characterization of the transcriptome at individual exon resolution. However, as a result of the massive amount of raw data, RNA-Seq poses new challenges to data analysis and interpretation. We present bioinformatic methods that enable effective analysis of RNA-Seq data either integrating or independent of known gene annotations. We demonstrate that such analyses can provide a more comprehensive profile of the mouse cardiac transcriptome. Indeed, a large number of novel transcripts and exons with potential biological significance were found from this study that had never been annotated previously. These findings open the way to further experimental investigations of their relevance in the pathogenesis of heart failure. Our study highlights the need to use both RNA-Seq and bioinformatic tools to achieve comprehensive evaluation of cardiac transcriptome in heart disease models as well as in human heart failure.

SUPPLEMENTAL MATERIAL

Analysis of Transcriptome Complexity via RNA-Seq in Normal and Failing Murine Hearts

Jae-Hyung Lee, Ph.D., Chen Gao, B.S., Guangdun Peng, Ph.D., Christopher Greer, B.S., Shuxun Ren, Ph.D., Yibin Wang, Ph.D., Xinshu Xiao, Ph.D.

From the Department of Integrative Biology and Physiology (J.H.L, G.P, C. Greer, X.X.), the Molecular Biology Institute (C. Gao, Y.W., X.X.); Departments of Anesthesiology (S.R., Y.W.), Physiology (S.R., Y.W.), Medicine (S.R., Y.W.), and the Cardiovascular Research Laboratories (C. Gao, S.R., Y.W.), David Geffen School of Medicine; University of California, Los Angeles.

Correspondence to Xinshu Xiao, Ph.D., 611 Charles E. Young Drive, Boyer Hall, Room 660, the Molecular Biology Institute, UCLA, Los Angeles, CA, 90095. Email: gxxiao@ucla.edu, Phone: 310-206-6522, Fax: 310-206-9184

Detailed Methods

RNA purification, library preparation and Illumina sequencing

Left ventricular tissues were collected from male C57BL/6 mice after 1 week (hypertrophy stage, HY) and 8 weeks post trans-aortic constriction (TAC) procedure (heart failure stage, HF), respectively, and their corresponding Sham controls (Sham-HY, Sham-HF). Doppler velocity measurements of right and left carotid arteries were obtained from TACed mice in order to ensure consistent pressure-gradient generated by the procedure. A 6 to 10 fold change in velocity ratio between right and left carotid arteries after TAC was required for a successful TAC. To conduct RNA-Seq analysis, total RNAs from six TAC and Sham-operated mice at the HY stage and four TAC and corresponding Sham mice at the HF stage were obtained. The hypertrophy and heart failure status of the TAC treated animals was established based on a significant increase in heart weight and a significant reduction in ejection fractions measured by echocardiogram [1-3]. Consistent with the literature [4], at 1-week post TAC, cardiac hypertrophy was detected while cardiac function was preserved. In contrast, at 8 weeks post-TAC, the mice demonstrated both hypertrophy and heart failure phenotypes (Online Table I).

Total RNA was isolated using the mirVana kit (Applied Biosystems), according to the manufacturer's instructions. We used the standard Illumina protocol to prepare libraries for RNA-Seq (http://www.illumina.com/support/documentation.ilmn). Briefly, 10μg total RNA was first processed via poly-A selection and fragmentation. We generated first-strand cDNA using random hexamer-primed reverse transcription and subsequently used it to generate second-strand cDNA using RNase H and DNA polymerase. Sequencing adapters were ligated using the Illumina Paired-End sample prep kit. Fragments of ~200 bp were isolated by gel electrophoresis, amplified by 15 cycles of PCR and sequenced on the Illumina Genome Analyzer II in the paired-end sequencing mode (2x72 bp or 2x76 bp reads).

Real-time PCR validation of gene expression

Total RNA was isolated from mouse Sham, hypertrophy and heart failure left ventricles using Trizol (Invitrogen) according to the manufacturer's protocol. For reverse transcription, 1ug of total RNA was used to generate first strand cDNA with Oligo-dT primer. Real-time PCR was performed using the SYBR Green Mix (Bio-Rad) on the CFX96 Real-time System (Bio-Rad).

RT-PCR validation of novel exons

cDNA libraries prepared from mouse Sham, hypertrophy and heart failure left ventricles were used to carry out PCR with primers targeting the two flanking constitutive exons of the novel exon. PCR was performed with Taq DNA polymerase (Invitrogen) using a C1000 Thermal Cycler (Bio-Rad). The sizes of the PCR products were confirmed using 4% Agarose gels. Subsequently, bands on the gel corresponding to the spliced inclusion form of the novel exons were purified using QIAquick Gel Extraction Kit (Qiagen) and sent to UCLA core facility for direct DNA sequencing.

Mapping of the RNA-Seq reads

In total, we obtained 168 million pairs of reads (2x72 or 2x76 bp in length) using the Illumina GA II sequencer. We mapped the reads to the reference genome and transcriptome using a two-step mapping strategy. In the first step, the reads were mapped to 49,039 Ensembl transcript sequences (release 56, September 2009) [5] using Bowtie [6] allowing up to 3 mismatches (~4% of total length). In the second step, we aligned the unmapped reads from the first step to the mouse genome (mm9) using BLAT [7] with the options: -stepSize=8, -fine. The output alignments were filtered to allow up to 5 mismatches, (~7% of total length). BLAT enabled identification of novel or known spliced junctions that connect two or more exons intervened by long introns. The above two steps were carried out for individual reads without considering the read-pairing information. Next, all read pairs were inspected to determine whether the mapped results support correct pairing of the reads according to the following criteria: 1) the pair of reads should be mapped to the same chromosome (study of trans-splicing across chromosomes is beyond the scope of this work); 2) the two reads of a pair should be mapped to opposite strands of the chromosome; 3) If one read in a pair is mapped to a gene, the other one should be mapped to the same gene or to an intergenic region. (Possible trans-splicing across Ensembl-annotated genes is ignored.) Since each read in a pair may map to multiple locations in the genome, all possible combinations of their mappings were examined for correct pairing. The pair of reads is considered as uniquely mapped if and only if one unique pair of mapped locations was identified. Altogether, about 95 million pairs of reads were uniquely mapped and used for further analyses (Table 1).

Expression measurement and analysis of differential gene expression

To measure expression levels of genes and exons, we used the variable RPKM (reads per kilobase of exon per million mapped reads) defined by Mortazavi et al. [8]. Analysis of differential gene expression was carried out for HY vs. Sham-HY or HF vs. Sham-HF, respectively. The number of uniquely mapped read-pairs for each gene in each sample was recorded. The total number of mapped reads in each lane was normalized using the upper-quartile normalization method [9]. Fisher's exact test was then performed using the above read counts for each gene. The resulted p-values were corrected via the Benjamini and Hochberg method as implemented in R. Finally, differentially expressed genes were defined as those with changes of at least 1.5-fold between a pair of samples at a false discovery rate (FDR) of 5% for genes expressed at $\geq$ 3 RPKM in $\geq$ 1 sample. Online Table III lists all genes differentially expressed between HY and Sham-HY or HF and Sham-HF.

Gene ontology (GO) enrichment analysis

The GO terms of each gene were obtained from the Ensembl database. To identify GO categories that are enriched in a specific set of genes (e.g., those that are differentially expressed in heart failure), the number of genes in the set with a particular GO term was compared to that in a control gene set. The control gene set was constructed so that the randomly picked controls and the test genes have one-to-one matched transcript length. Based on 10,000 randomly selected control sets, a p-value for enrichment of each GO category in the test gene set was calculated as the fraction of times that F_{test} was lower than or equal to $F_{control}$, where F_{test} and $F_{control}$ denote, respectively, the fraction of genes in the test set or a random control set associated with the current GO category. A p-value cutoff (1/total number of GO terms considered) was applied to choose significantly enriched GO terms.

Analysis of differential isoform expression and dissimilarity

To estimate the expression level of each isoform of the Ensembl genes, we used the Cufflinks software (v0.9.2) [10] where the gene structures were defined according to Ensembl v56 annotations. We used the "cuffdiff" module to estimates isoform expression levels of each gene and test whether the isoforms are differentially expressed between two samples. Based on the outputs from "cuffdiff", we identified genes with significant isoform expression changes ($p < 0.05$) due to alternative transcription start site (ATSS) or alternative splicing (AS) comparing HY and Sham-HY or HF and Sham-HF (Figure 2A and Online Table IV). To further assess overall isoform expression dissimilarity in two samples (A and B), a dissimilarity score was defined based on the Morisita-Horn similarity index as follows:

$$dissimilary\ score\ = 1 - \frac{2\sum_i P_i(A)P_i(B)}{\sum_i [P_i^2(A) + P_i^2(B)]}$$

where $P_i(A)$ and $P_i(B)$ represent the expression of isoform i normalized by overall gene expression in the sample A or B. We only considered highly expressed genes (RPKM $\geq$ 3) in this analysis. A total of 265 genes had dissimilarity scores differing by more than 0.2 between two different heart failure stages (Figure 2C).

Guided transcriptome reconstruction

We developed a method to reconstruct alternative isoforms and discover novel isoforms of known genes. In the first step, we defined novel expressed sequence fragments (seq-frags) in each gene by 1) identifying the expressed blocks which are connected to known exons by junction reads or 2) extending expressed regions from the known exon boundaries. To define a seq-frag, we imposed a minimum expression level cutoff to control for background noise. For each gene, the expression levels in the intronic regions were fitted by a binomial distribution and the expression level cutoff was determined as the value at the 95% of the cumulative distribution. Another expression level cutoff was defined as 10% of the average expression of all Ensembl exons in the gene. The smaller value of the two cutoffs was chosen as the final cutoff to define novel seq-frags. To reduce possible false positive isoforms, we required at least 2 spliced junction reads originating from each end of a seq-frag.

We then constructed a directed graph in which known exons or novel seq-frags defined above were considered as nodes and their connections represented by spliced junction reads as edges of the graph. All possible paths of the graph were then determined originating from the root node (the 5'-most known exon or seq-frag of the gene). Note that isoforms constructed in this way may contain false positives since not all possible paths of the graph may correspond to an actual expressed isoform. Further filtering may be applied to reduce false positives, such as by calculating the likelihood of a path being an authentic isoform using the pairing information of the paired-end reads. However, since our main goal is to identify as many novel *exons* as possible, we did not implement such filtering steps in order to avoid loss of authentic novel exons.

Analysis of the features of novel exons

To calculate the conservation levels of novel exons, we used the PhastCons scores for the multiple alignments of 30 vertebrate genomes (with mouse as the reference) in the UCSC genome database [11, 12]. The conservation scores for individual exons or introns were calculated by taking the average PhastCons score per base in the region. The 5' and 3' splice site strength were evaluated using the MaxEnt method [13]. Three features (splice site strength, expression level and exon length) were included in the principal component analysis (PCA). In Figure 3C, the representative ellipses for novel skipped exon and known skipped exon were obtained by fitting the principal components with a Gaussian distribution. The central points and rotation angles of the ellipses were calculated based on the average and standard deviation of the fitted distributions.

Prediction of miRNA target sites in alternative termination exons

We downloaded the 46-way multiz alignments from the UCSC browser and prepared the multiple sequence alignments for 23 species for the alternative termination exons as needed by the TargetScan programs [14, 15]. Targetscan predicted the miRNA target sites by considering the context features and conservation levels of the putative target sites in different species. A total of 892 alternative termination exons had predicted miRNA target sites for which the context score is less than -0.2 or the target site is conserved, when all known mouse miRNAs were included in the analysis. Furthermore, we collected 102 miRNAs that are known to be expressed in mouse heart [16-18]. About 39% of the alternative termination exons had predicted target sites for the miRNAs expressed in heart.

RNA-Seq data of other mouse tissues

To validate the expression of novel exons or novel transcript clusters (NTCs) in other tissues or cell types, we collected publically available RNA-Seq datasets (for 12 tissues/cell types) from the NCBI Short Read Archive (SRA) (IDs: SRP000198, SRA021498, SRA012498). All reads were preprocessed and mapped to the reference sequences as described in the above section for read mapping. Gene and exon expression levels were analyzed in the same way as described above.

Identification of potential NMD targets and protein domain analysis

To investigate the impact of novel exons on gene expression, we focused on questions including whether in-frame premature termination codons (PTCs) are introduced, whether nonsense-mediated decay (NMD) of the mRNA may be activated or truncated/defective proteins may be generated. Briefly, the novel exons were introduced into all known Ensembl transcripts of the corresponding gene to determine whether a PTC can be generated. If a PTC resides more than 55 nucleotides upstream of the 3'-most exon-exon junction, the gene or transcript is classified as a potential target of the NMD pathway. For protein domain analysis, we predicted functional domains in translated amino acid sequences of all transcripts using the Pfam server [19]. We examined the protein domain organizations prior to and following introduction of a novel exon into the transcript.

Construction of NTCs in intergenic regions (*de novo* reconstruction)

To identify NTCs, all reads mapped to intergenic regions were collected and sorted according to chromosomes. Since we only focused on multi-exon NTCs in this study, we first inspected spliced junction reads to find evidence of splicing. Only junction reads corresponding to the canonical splice site signals (GT-AG, GC-AG and AT-AC) in either strand were included for subsequent analyses. Extending from the identified exon junctions, we searched for expressed seq-frags by checking for mapped reads in the neighborhood. The boundaries of expressed seq-frags were defined if 1) another exon junction is observed, or 2) expression level tapered off. Seq-frag definition and isoform reconstruction were conducted similarly as in the guided transcriptome reconstruction. Finally, the novel transcript isoforms were clustered together to generate NTCs.

The NTCs were classified into coding and non-coding clusters using the Coding Potential Calculator (CPC) software [20]. This method assesses coding capability using the support vector machine algorithm. We evaluated all isoforms in a NTC and classified the NTC as coding if at least one transcript isoform is predicted as coding.

Tissue specificity index

To characterize the tissue-specific expression of NTCs and long non-coding RNAs (lncRNAs), we defined a tissue specificity index for each gene, $T(G)$, as follows:

$$T(G) = 1 - H(G)$$

$$H(G) = -\sum_{t=1}^{n} p(g_t)\log_2 p(g_t)$$

where n is the total number of tissues and $p(g_t)$ represents the expression level (RPKM) of gene g normalized by the sum of expression levels of all genes in tissue t. $H(G)$ was normalized by the total number of tissues (i.e., divided by $\log_2 n$) so that its value is between 0 and 1. In this definition, larger $T(G)$ values correspond to more significant tissue specificity.

Comparison of novel exons and NTCs to the proteomics database

To check whether novel exons or NTCs could be translated or not, we compared the amino acid sequences of novel exons and NTCs with known proteomics data in the "PRIDE" database [21]. This database is a centralized, standards compliant, public data repository for proteomics data. For both novel exons and NTCs, we performed Blastx of the translated novel sequences against expressed protein sequences in the database with an E-value cutoff of 1e-10. After getting all alignments between the novel exons or transcripts and protein sequences, we calculated the overall identity percentage (%id) for the alignments. Since the novel sequences may have sequence similarity to known protein-coding genes, the presence of alignments passing the E-value cutoff may not suggest that they are protein-coding. To resolve this problem, we carried out another round of Blastx of the novel sequences against known protein sequences (combining UCSC genes, Refseq and ensembl v61 proteins). We considered the novel sequences as putatively protein-coding only if its %id against the PRIDE database was greater than that against the known protein sequences. In total, we identified 174 novel exons and 48 NTCs (out of the 105 predicted as coding) that may code for proteins. Among the putatively translated novel exons and NTCs, 40 and 27, respectively, had expression level difference of at least 2-fold in the HF or HY (or both) sample compared to the Sham controls.

Co-expression network analysis

In total, 10,091 known genes and 98 novel transcript clusters (required to have RPKM≥3) were included in the co-expression network analysis. The expression levels of each gene in 8 samples (2 technical replicates for HY, HF and each Sham sample) were calculated. Network construction and module detection were performed using the WGCNA package implemented in R [22]. Following recommended procedures and parameters to threshold the hierarchical clustering tree, we detected 52 modules in total, among which 20 (Online Table XI) were enriched with differentially expressed genes between HY and Sham-HY or HF and Sham-HF. We thus focused on the 20 modules in the functional analysis. In each module, the 10 most highly connected NTCs were identified, each of which was required to have at least 10 neighboring known genes. In this way, 15 NTCs were chosen and their neighboring genes were subject to GO analysis (Online Table XII).

Graphical representations and bioinformatic analysis

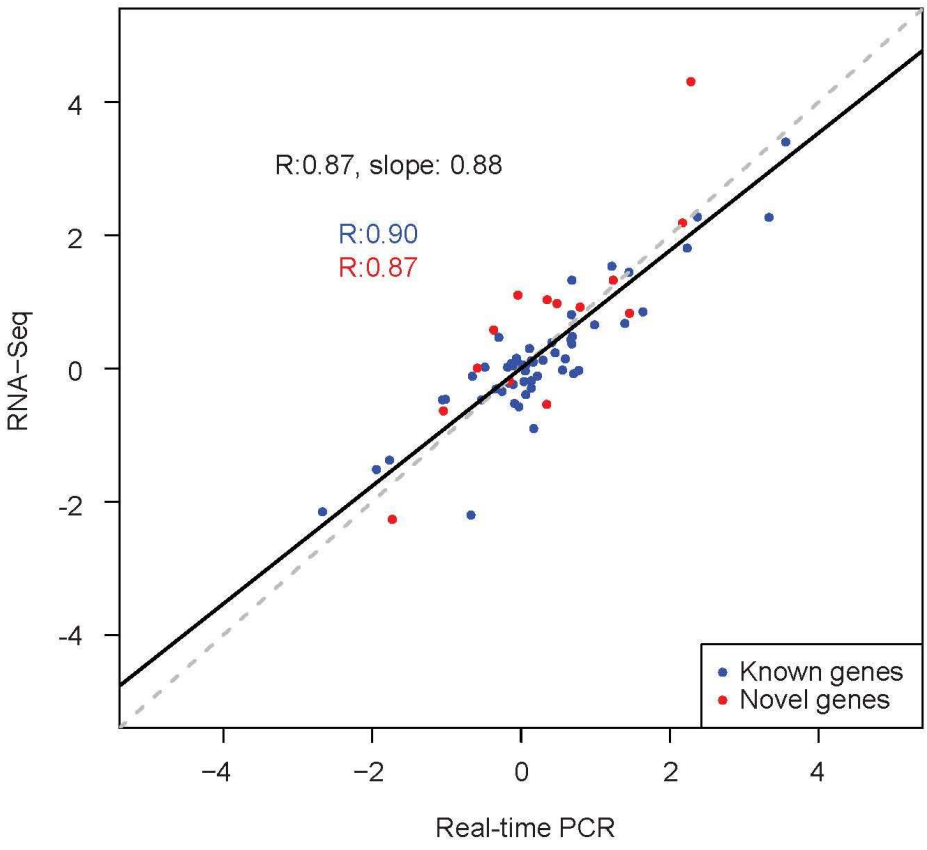
The proportional Venn diagrams were produced via the BioVenn website [23]. Read distribution plots and cumulative distribution functions were generated using custom scripts in R. Heatmaps (Figures 3D and 5A) were drawn using the MeV software package [24]. Network representations in Figure 6A was generated using Cytoscape [25]. All bioinformatic analyses including statistical tests were performed using in-house scripts written in Python, Perl or R.

Supplemental Results

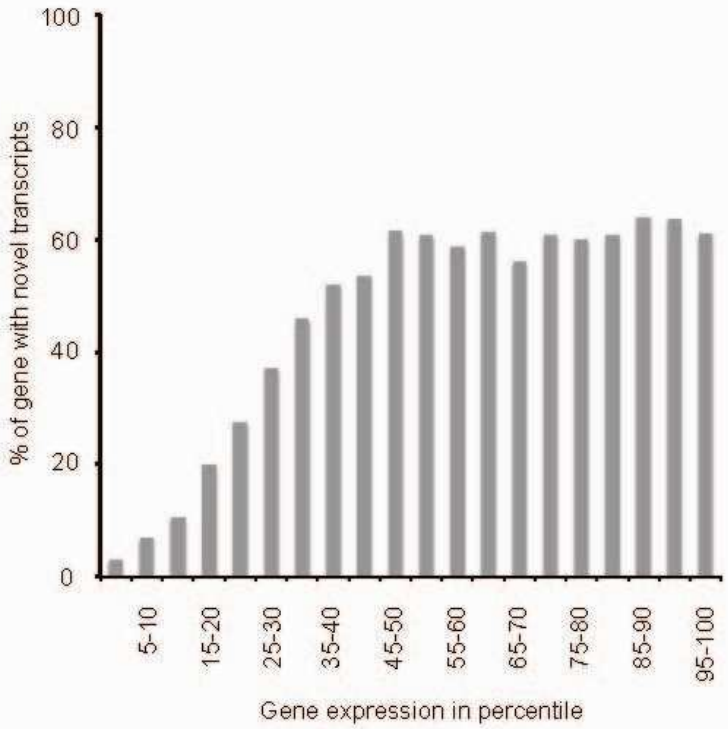
Impact of novel exons on gene expression

To investigate the impact of novel exons on gene expression, we focused on questions including whether in-frame premature termination codons (PTCs) are introduced, whether nonsense-mediated decay (NMD) of the mRNA may be activated or truncated/defective proteins may be generated. We found that

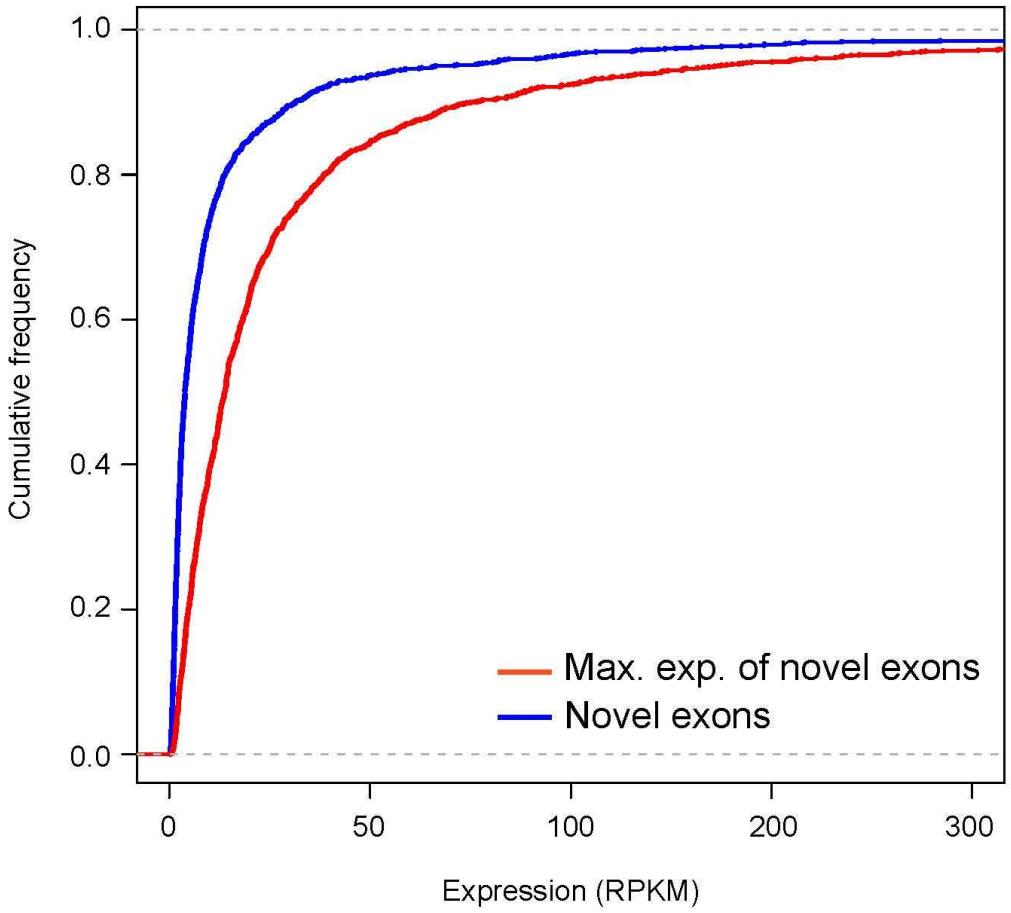
825 novel exons in 558 genes can generate PTC-harboring transcripts. Among the 392 exons that do not lead to introduction of PTCs, 201 were located in the UTR regions and none of those in the coding regions affected known protein domains as annotated in the Pfam database [19]. Thus, there might exist selection pressure to protect a gene's protein-coding potential from being disrupted by these exons. On the other hand, 117 (14%) of the PTC-inducing novel exons disrupted the predicted protein domain organizations of the gene, mostly by generating truncated proteins. Some of these genes are known to have important functions in the heart, including cytoskeleton-related genes (*Myo19, Myh7b, Snapc4, Spna1* and *Spnb2*), kinase genes (*Mark4, Blk, Prkca* and *Ripk1*) and RNA binding proteins (*Tia1, Myef2* and *Spen*). Interestingly, the majority (95 out of 117) of exons with the potential to generate truncated proteins are predicted to trigger the NMD pathway [26, 27]. In contrast, 468 (57%) of all the PTC-inducing novel exons can activate this pathway. Thus, as a surveillance mechanism, NMD is preferably used to prevent production of defective protein products.



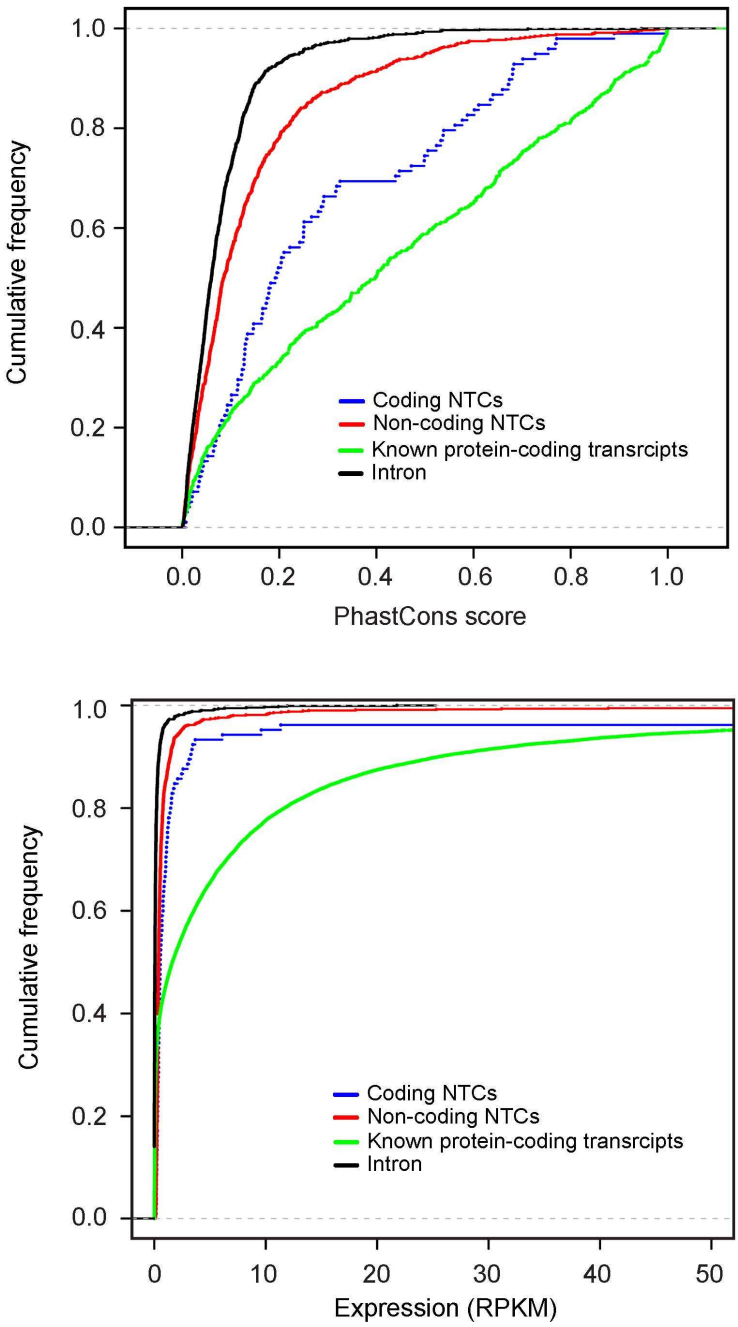
Online Figure I. Real-time PCR validation of gene expression levels estimated by RNA-Seq. X-axis: ΔΔCt values from real-time PCR comparing HY or HF and their corresponding sham controls normalized by the expression level of *Gapdh*. Y-axis: $\log_2$(fold-change) between HY or HF and their corresponding sham controls estimated via RNA-Seq. Pearson correlation coefficient (R) based on all genes is shown in black. Blue dots and text correspond to known Ensembl genes, red dots and text correspond to novel transcript clusters identified in our study.



Online Figure II. Distributions of novel isoforms as a function of gene expression levels. Percent of Ensembl genes with novel isoforms detected by the guided transcriptome reconstruction. Genes were ranked according to their expression levels (low to high, x-axis: percentile of the gene in the ranking of all genes). To show the entire spectrum of gene expression levels, we included all genes expressed above 0 RPKM.



Online Figure III. Expression levels of novel exons. Cumulative distribution functions of expression levels of novel exons in the heart samples in our study (blue) and the maximum expression (max. exp.) level of the novel exons (red) across all available RNA-Seq data as shown in Figure 3D.



Online Figure IV. Conservation and expression levels of novel transcript clusters. Cumulative distribution functions of conservation and expression levels of coding or non-coding NTCs in our study, known protein-coding genes, and random intronic regions.

Online Tables

Online Table I. Phenotypes of 1-week and 8-week post-TAC mice and corresponding Sham controls. BW: body weight, HW: heart weight, EF: ejection fraction, HR: heart rate.

Group	BW (g)	HW (mg)	HW/BW	%EF	HR (bpm)
TAC (1 week)	30.9 ± 0.58	147 ± 14.6	4.74 ± 0.27	55.39 ± 3.90	506.3 ± 55.1
Sham	32.8 ± 1.43	138 ±1.1	4.22 ± 0.22	55.35 ± 6.81	532.0 ± 47.4
***t*-test**	$p = 0.06$	$\mathbf{p = 0.008}$	$\mathbf{p = 0.025}$	$p = 0.99$	$p = 0.51$
Group	**BW (g)**	**HW (mg)**	**HW/BW**	**%EF**	**HR (bpm)**
TAC (8 week)	31.6 ± 1.47	245 ± 22.4	7.77 ± 1.09	26.59 ± 6.18	489.3 ± 41.9
Sham	33.5 ± 2.21	172 ± 3.3	5.13 ± 0.91	52.61 ± 3.72	467.7 ± 32.5
***t*-test**	$p = 0.23$	$\mathbf{p = 0.017}$	$\mathbf{p = 0.019}$	$\mathbf{p = 0.00036}$	$p = 0.52$

Online Table II. Primer sequences used for validation of expression of known genes and novel genes.

Gene	Primer1	Primer2
A2bp1	GTGGTTATGCTGCGTACCG	GGAGCAAGTGTGTGGTGGTA
Actb	AGCCATGTACGTAGCCATCC	GCTGTGGTGGTGAAGCTGTA
Angptl4	ACGCTTATGAGCTACGGGCTCCA	GCCTTGGGTGCAGCAACGCT
Cfd	GTGGAACCCGGCACGCTCTG	TGCCGGAGTCTCCCCTGCAA
Clip2	TACGGCACGGATCGTTTCGCG	AGCGGTCACAAAGGCCGTGT
Col6a1	CCGGCGCAATTTCACGGCAG	TCCTCTGGCAGCCTGGCACT
Cugbp1	GTTCCAAGGACCTGGTCTGA	CCAGGGAGGACCTTCATGTT
Cugbp2	TCGGAAAAGGAGCTGAAAGA	GCCTCAAGTGCAGCTTTTCT
Ddx23	CAAGGACCGAAAAAGATCCA	GTCGCTTGTCTCCATGTTCA
Egr1	AGCGAACAACCCTATGAGCACC	ATGGGAGGCAACCGAGTCGTTT
Fos	CACTGCCCGAGCTGGTGCATT	ACACAGACCAGGCCTTGACTCAC
Hnrnpa1	TGGAAGCAATTTTGGAGGTGG	GGTTCCGTGGTTTAGCAAAGT
Hnrnpa2b1	AAGAAATGCAGGAAGTCCAAAGT	CTC CTCCATAACCAGGGCTAC
Hnrnpc	CCTCCTCCTCCTCCTATTGC	TTGGAAGAAGATCCCCTTTG
Hnrnpd	GATCCTAAAAGGGCCAAAGC	ACCTCACCAAAACCACCAAA
Hnrnpf	GGCATCTGTGGTGGTTCTTT	TGCAGTCGGAGAGGAAGTTT
Hnrnpu	GTCTCCTCAGCCACCTGTTG	TTGCCTTTTGACACACCGTA
Hnrpdl	CAAACTGGATGGCAAATTGA	CTCTCCAAAGGCTCCAAAAT
Inmt	ACTACAGCTTCCACTCTGGCCCT	GCCCCCTACACCTCCTGTAGAG
Junb	CCGAGAAAGCAGGCGCACCA	GCGTCCTCGGGAGCTGGAGA
Lsm7	ACCATTCGGGTGAAGTTCCAGG	GAGTGTCCTCCGTCAGCTTGTA
Mbnl2	GCCGCTGTTCAAGAGAGAAC	GGCGTTCCTGGAAACATAAA
Mrpl42	CGCGGTGAGTAGCCGTGGAG	GGCGAGGTCCACCTGGCAAT

Mybpc2	GCCACTGCGGTCCCAAGTCC	AGGTGGGGCCTCTTTGGGGG
Myh7	TGAGGAGGCGGAGGAGCAGG	CACGGGCACCCTTGGAGCTG
Nppb	AGGGAGAACACGGCATCATT	GACAGCACCTTCAGGAGAT
Ptbp1	AGCAGAGACTACACTCGACCT	GCTCCTGCATACGGAGAGG
Rbm17	TCTGCAGGGGAAGTTCTGAT	GGGTCTGGTCGTCTTGAAAA
Rbm9	AACCAGGAGCCAACAACAAC	ACTCCCGTAGAGGGTCAGGT
Rbmx	ACTTTTGAAAGCCCAGCAGA	GGTTCCTCCACTTCCTCCTC
Rnps1	GAAGAGACGCAGTGCTTCAA	CGGAGGAGCTGCTAGAACTG
Rtp4	AGCAGACAGTGCTTGGCAGGTTC	CGTGCCTGGCCCTGCGATTT
Scn4b	GGGCACTGGGCTTTTGGGTCT	GGTGCAGGGCAGCAGGATCG
Sfrs1	CACTGGTGTCGTGGAGTTTG	GGCTTCTGCTACGACTACGG
Sfrs3	TGGAACTGTCGAATGGTGAA	GACGCTGAAAGGGCTAGTTG
Sfrs4	AAAAGCTGGACGGAACTGAA	CTCTTCGAATGGCTGCTTTT
Sfrs5	TGCTCCACCTGTAAGAACAGAA	GTGCATCCGCAAAGGTTACT
Sfrs7	ATCGCTATAGCCGACGAAGA	CGAGGAGATGCTGATCTTGA
Tmem82	GGGGCCCTTGGAGTCTCGGT	TCCCGGAACCCTTGCCCAGA
U2af1	GAGATGAACGTCTGCGACAA	ATCGGCTGTCCATTAAACCA
U2af2	ACTCCTGATGGTCTGGCTGT	TCAGTAATGCCAAAAGGGATG
Vim	CCAGAGACCCCAGCGCTCCT	GCCGGAGCCACCGAACATCC
chr11:62416331-62418311+	CTGGGGTGAGGAGTGGTCTA	AGCTCAACGAGCGAGAAGAG
chr8:94498099-94515675-	GATTTCCAGCCTGCTTCTTG	CGGTTGTGTGAGCTCCTTTT
chr19:28845241-28862777+	GAGTCTGTTCAAGGGGATCG	CGTCCAGTGAAACTGCAAAA
chr1:188411121-188439353+	GTGGCTGCTGAGGTCAGACT	AACAGGAAACGTGGAGTCGT
chr6:134879119-134901898+	TGTGGAAGGGGGAGATACAG	GAGATGCAGTCCGGTAGAGC
chr17:35133686-35137288-	CCTCTGGAGAAAAGGTGCTG	GAGTATGACCAGGCCCAAGA
chr13:103119664-103127036+	GAAGTGCATTGAGGCTGTGA	CCCATTTGACGACATGATGA

Online Table III. List of differentially expressed genes. All differentially expressed genes between HY or HF and the corresponding shams were listed with fold-change values. "-" denotes not differentially expressed. UP: up-regulated; DN: down-regulated.
Online Table III is supplied as an Excel datasheet.

Online Table IV. List of genes with differential isoform expression due to alternative splicing (AS) or alternative transcription start sites (ATSS). N: not differentially expressed; Y: differentially expressed.
Online Table IV is supplied as an Excel datasheet.

Online Table V. GO categories of genes with differential isoform expression due to alternative transcription start sites (ATSS) or alternative splicing (AS). Only genes whose differential expressions are common to the HY and HF stages were analyzed. Red color: related to muscle function,

blue color: related to ATP synthesis process.

GO ID	P-value	Description
GO terms enriched in genes with ATSS difference		
GO:0015986	< 0.0001	ATP synthesis coupled proton transport
GO:0045263	< 0.0001	proton-transporting ATP synthase complex, coupling factor F(o)
GO:0000276	< 0.0001	mitochondrial proton-transporting ATP synthase complex, coupling factor F(o)
GO:0005739	< 0.0001	mitochondrion
GO:0055010	< 0.0001	ventricular cardiac muscle morphogenesis
GO:0055003	< 0.0001	cardiac myofibril assembly
GO:0060048	< 0.0001	cardiac muscle contraction
GO:0016459	0.0003	myosin complex
GO:0060047	0.0006	heart contraction
GO:0045214	0.0008	sarcomere organization
GO:0015986	0.0001	ATP synthesis coupled proton transport
GO terms enriched in genes with AS difference		
GO:0008746	< 0.0001	NAD(P) transhydrogenase activity
GO:0032781	< 0.0001	positive regulation of ATPase activity
GO:0014853	< 0.0001	regulation of excitatory postsynaptic membrane potential involved in skeletal muscle contraction
GO:0030018	< 0.0001	Z disc
GO:0030017	< 0.0001	sarcomere
GO:0005862	< 0.0001	muscle thin filament tropomyosin
GO:0045214	0.0004	sarcomere organization
GO:0008016	0.0019	regulation of heart contraction
GO:0005861	0.002	troponin complex
GO:0055010	0.0021	ventricular cardiac muscle morphogenesis

Online Table VI. List of novel alternative splicing events identified in Ensembl genes. Genomic coordinates shown according to the mm9 mouse genome assembly. SE: skipped exon; RI: retained intron; A3E: alternative 3'ss exon; A5E: alternative 5'ss exon; MXE: mutually exclusive exon.
Online Table VI is supplied as an Excel datasheet.

Online Table VII. Primer sequences used for validation of novel alternatively spliced exons. "Y" in confirmation column means expression of the novel exon was confirmed. Genomic coordinates shown according to the mm9 mouse genome assembly.

Ensembl ID	Gene symbol	Confir-mation	Novel exon coordinate	Primer1	Primer2
ENSMUSG00000031575	Ash2l	Y	chr8:26940847:-_chr8:26940864:-	CCCAAGTTTGGACTTTTGGA	CCAAAGGATAGCCATGAGGA
ENSMUSG00000019943	Atp2b1	Y	chr10:98481912:+_chr10:98481998:+	GTGGCCAGATCTTGTGGTTT	GGAGGGTTGTCGCCTTAGAG

ENSMUSG00000021820	Camk2g	Y	chr14:21560600:-_chr14:21560713:-	GGTCTACGGTGGCATCCAT	TAGGTGCACGCTGGAACTC
ENSMUSG00000068823	Csde1	Y	chr3:102840498:+_chr3:102840644:+	TCCTTTGGAACTTGTGCTGA	AACCCCAGTTTCACGAAGTG
ENSMUSG00000022475	Hdac7	Y	chr15:97631505:-_chr15:97631760:-	GGACGTTTGTGATGCTACCC	TGCACCACTGTCTTCCACTC
ENSMUSG00000019726	Lyst	Y	chr13:13779100:+_chr13:13779255:+	GAGTTCTCCAGGCTGCTTTG	TCTTGGCTCATCCTCCTCTG
ENSMUSG00000001419	Mef2d	Y	chr3:87947272:+_chr3:87947421:+	CTTCGCGTAACCGAGGATT	AGGAGCCTCACACTGTGCTC
ENSMUSG00000062991	Nrg1	Y	chr8:32929525:-_chr8:32929666:-	CCAGAGAAACCCCTGACTCC	GGCCTCTCTTCTTCCACAAA
ENSMUSG00000021488	Nsd1	Y	chr13:55413307:+_chr13:55413344:+	GCGAAGTCACCAAGGAAAGA	CGAGCTTGGAAATGAAAAGC
ENSMUSG00000019907	Ppp1r12a	Y	chr10:107712814:+_chr10:107712944:+	TGAAATGGAAGAAGAGCTCAAA	AAGTCCTGCTGCTTTGCTTC
ENSMUSG00000019907	Ppp1r12a	Y	chr10:107686012:+_chr10:107686125:+	AAAGCCCATGGCTTCTGTAA	GCACCGTAACTGCCAGTCTT
ENSMUSG00000032827	Ppp1r9a	Y	chr6:5033041:+_chr6:5033106:+	TGAAATGGAAGAAGAGCTCAAA	TGACAGTCTCCCGTGAGTTC
ENSMUSG00000050965	Prkca	Y	chr11:107806871:-_chr11:107806918:-	AGAGCATGCCTTCTTCAGGA	GGGAGGCATTTGATCTTTCA
ENSMUSG00000038132	Rbm24	Y	chr13:46517645:+_chr13:46517757:+	AGCTTCACCCAGCCCTTATC	ATGCTGCATAGGGGTACTGG
ENSMUSG00000031808	Slc27a1	Y	chr8:74094095:+_chr8:74094211:+	TTCTCGGAGTCTGGAATGCT	TGCAGACGATACGCAGAAAG
ENSMUSG00000032114	Slc37a4	Y	chr9:44209790:+_chr9:44209852:+	GCAGCATCCATGTTCCTCTT	ATTGGCCATAAGTCCCACAA
ENSMUSG00000038485	Socs7	Y	chr11:97238809:+_chr11:97239000:+	ACGCTGCCTACATCTGTCCT	CTACTTCAGGGGCTTCACCA
ENSMUSG00000044308	Ubr3	Y	chr2:69809459:+_chr2:69809491:+	GCTCTGACGCTTTGGACTCT	TGTGGTTAACTGGGGAGAGG
ENSMUSG00000033909	Usp36	Y	chr11:118147968:-_chr11:118148085:-	CTCTACACTTGGCGGTGTTG	GAAGCACCTTCTTGGCAGAG
ENSMUSG00000024095	Hnrpll	Y	chr17:80444280:-_chr17:80444362:-	ATATGCAAGGCCAACTCGTC	TAAAGGCAGTAGCGGACCAT

ENSMUSG00000024384	Iws1	Y	chr18:32253193:+_chr18:32253351:+	ACCTGGAGATCCTGGCTTCT	TCGATGCTGATTTTCACTGC
ENSMUSG00000053470	Jmjd1a	Y	chr6:71566812:-_chr6:71566864:-	TGGAGCTGTAAAACGCAAGT	CATCCTTGAGGAAGTTTTGCT
ENSMUSG00000038056	Mll3	Y	chr5:24836194:-_chr5:24836241:-	CCTGGAACTCAGGGTACTGC	TGCTCCATCTGTGACCATTT
ENSMUSG00000020527	Myo19	Y	chr11:84720256:+_chr11:84720373:+	CTGAGAAGGCTCGGACCTC	GGCCTCAGACCACAGTTAGC
ENSMUSG00000025060	Slk	Y	chr19:47715345:+_chr19:47715419:+	GAAAAATGCCACCTGTTGGT	CGGACTCCCCAGTCATTTTA
ENSMUSG00000032826	Ank2	Y	chr3:126689503:-_chr3:126689556:-	CCTGAGACAATGACGGAGGT	GGATCTTCCTCCCTCCAGAC
ENSMUSG00000032826	Ank2	Y	chr3:126673921:-_chr3:126673956:-	CGTGGCTCTTTCTTCCAGTC	GAAACTGAGCTCCCGAAGGT
ENSMUSG00000032826	Ank2	Y	chr3:126629721:-_chr3:126629812:-	CCTTCCTTCAGCCAAAGATG	AGTGGGTTTGTGGTTTCTGG
ENSMUSG00000054808	Actn4	N	chr7:29680140:-_chr7:29680205:-	CAATGGAGCATATCCGTGTG	GTCACGAGGCCACTATGGTT

Online Table VIII. Novel alternative terminal exons identified in Ensembl genes and categorization of different types of exons. See main text for the definitions of the types of novel terminal exons. "Original": terminal exons identified relative to Ensembl v56. "Updated": terminal exons identified relative to the most recent databases including Ensembl v61, UCSC KnownGenes and RefSeq genes.

Types of novel terminal exons	5' type (1)	5' type (2)	3' type (1)	3' type (2)
Original terminal exons	963	129	1,684	57
% of total (2,833)	34%	5%	59%	2%
Updated terminal exons	757	91	1,554	51
% of total (2,453)	31%	4%	63%	2%

Online Table IX. List of novel alternative terminal exons. Genomic coordinates shown according to the mm9 mouse genome assembly. See main text for the definition of different types of exons.
Online Table IX is supplied as an Excel datasheet.

Online Table X. List of novel transcript clusters. All novel transcript clusters were listed with expression level (RPKM) in different samples (HY, Sham-HY, HF and Sham-HF), coding potential (Y: coding, N: non-coding).
Online Table X is supplied as an Excel datasheet.

Online Table XI. Twenty significant modules in the co-expression network enriched with differentially expressed genes. Number of known genes (KG), novel transcript clusters (NTC), and differentially expressed (DG) genes are listed with enriched GO terms.

module	KG	NTC	DG	enriched	GO ID	P-value	description
salmon	160	0	50	Sham-HF	GO:0007264	0.0003	small GTPase mediated signal transduction
					GO:0005743	0.0004	mitochondrial inner membrane
					GO:0050873	0.0007	brown fat cell differentiation
					GO:0030286	0.0013	dynein complex
					GO:0007264	0.0003	small GTPase mediated signal transduction
midnightblue	149	3	52	Sham-HF	GO:0016790	< 0.0001	thiolester hydrolase activity
					GO:0005739	< 0.0001	mitochondrion
					GO:0016874	0.0005	ligase activity
					GO:0005759	0.001	mitochondrial matrix
darkgrey	63	2	33	Sham-HF	GO:0005739	0.0001	mitochondrion
					GO:0009055	0.0002	electron carrier activity
					GO:0055114	0.0007	oxidation reduction
					GO:0016491	0.0011	oxidoreductase activity
					GO:0022900	0.0013	electron transport chain
					GO:0000166	0.0031	nucleotide binding
lightcyan	149	2	58	Sham-HF	GO:0006810	< 0.0001	transport
					GO:0005743	< 0.0001	mitochondrial inner membrane
					GO:0005739	< 0.0001	mitochondrion
					GO:0070469	0.0001	respiratory chain
					GO:0022900	0.0001	electron transport chain
					GO:0051291	0.0004	protein heterooligomerization
					GO:0005759	0.0009	mitochondrial matrix
purple	562	10	164	HF	GO:0008360	< 0.0001	regulation of cell shape
					GO:0017124	< 0.0001	SH3 domain binding
					GO:0070059	< 0.0001	apoptosis in response to endoplasmic reticulum stress
					GO:0007015	< 0.0001	actin filament organization
					GO:0005515	< 0.0001	protein binding

					GO:0005737	< 0.0001	cytoplasm
					GO:0003779	< 0.0001	actin binding
saddlebrown	37	0	8	Sham-HF	GO:0005739	< 0.0001	mitochondrion
					GO:0055114	0.0011	oxidation reduction
					GO:0008152	0.0015	metabolic process
					GO:0016491	0.0027	oxidoreductase activity
darkturquoise	77	1	32	Sham-HF	GO:0008152	< 0.0001	metabolic process
					GO:0005739	< 0.0001	mitochondrion
					GO:0005759	0.0001	mitochondrial matrix
					GO:0003824	0.0005	catalytic activity
					GO:0005743	0.0012	mitochondrial inner membrane
					GO:0031072	0.002	heat shock protein binding
darkred	87	1	17	Sham-HF	GO:0005739	0.0002	mitochondrion
					GO:0000166	0.0003	nucleotide binding
					GO:0007179	0.0015	transforming growth factor beta receptor signaling pathway
red	675	10	271	HY, HF	GO:0005587	< 0.0001	collagen type IV
					GO:0005515	< 0.0001	protein binding
					GO:0005737	< 0.0001	cytoplasm
					GO:0043277	< 0.0001	apoptotic cell clearance
					GO:0005739	0.0001	mitochondrion
					GO:0016020	0.0002	membrane
					GO:0005856	0.0003	cytoskeleton
brown	745	10	419	HY, HF, Sham-HF	GO:0000120	< 0.0001	RNA polymerase I transcription factor complex
					GO:0050840	< 0.0001	extracellular matrix binding
					GO:0001568	< 0.0001	blood vessel development
					GO:0005604	< 0.0001	basement membrane
					GO:0051301	< 0.0001	cell division
					GO:0005581	< 0.0001	collagen
					GO:0005515	< 0.0001	protein binding
cyan	157	0	50	Sham-HF	GO:0005743	< 0.0001	mitochondrial inner membrane
					GO:0003824	< 0.0001	catalytic activity
					GO:0005739	< 0.0001	mitochondrion
					GO:0008016	0.0001	regulation of heart contraction
					GO:0046716	0.0002	muscle maintenance
					GO:0070469	0.0005	respiratory chain
					GO:0007050	0.0006	cell cycle arrest
mediumpu	27	1	2	HY	GO:0045121	0.0004	membrane raft

rple3							
darkorange	49	0	8	HY, Sham-HY	GO:0006464	0.0029	protein modification process
					GO:0008134	0.0031	transcription factor binding
grey60	148	0	62	Sham-HF	GO:0005743	< 0.0001	mitochondrial inner membrane
					GO:0005759	< 0.0001	mitochondrial matrix
					GO:0005739	< 0.0001	mitochondrion
					GO:0042645	0.0002	mitochondrial nucleoid
					GO:0046872	0.0003	metal ion binding
					GO:0008270	0.0008	zinc ion binding
					GO:0006915	0.001	apoptosis
turquoise	956	8	9	Sham-HY	GO:0006974	< 0.0001	response to DNA damage stimulus
					GO:0030529	< 0.0001	ribonucleoprotein complex
					GO:0003723	< 0.0001	RNA binding
					GO:0003674	< 0.0001	molecular_function
					GO:0005515	< 0.0001	protein binding
					GO:0042765	< 0.0001	GPI-anchor transamidase complex
					GO:0005737	< 0.0001	cytoplasm
greenyellow	500	3	72	Sham-HF	GO:0006810	< 0.0001	transport
					GO:0005515	< 0.0001	protein binding
					GO:0005743	< 0.0001	mitochondrial inner membrane
					GO:0005739	< 0.0001	mitochondrion
					GO:0015909	< 0.0001	long-chain fatty acid transport
					GO:0022900	< 0.0001	electron transport chain
					GO:0004129	< 0.0001	cytochrome-c oxidase activity
lightgreen	141	3	44	Sham-HF	GO:0005743	< 0.0001	mitochondrial inner membrane
					GO:0005739	< 0.0001	mitochondrion
					GO:0032981	0.0001	mitochondrial respiratory chain complex I assembly
					GO:0000287	0.0002	magnesium ion binding
					GO:0022900	0.0002	electron transport chain
					GO:0070469	0.0009	respiratory chain
tan	246	3	84	Sham-HF	GO:0005739	< 0.0001	mitochondrion
					GO:0016772	0.0001	transferase activity,

							transferring phosphorus-containing groups
					GO:0007031	0.0002	peroxisome organization
					GO:0003823	0.0003	antigen binding
					GO:0042593	0.001	glucose homeostasis
sienna3	30	1	5	HY, Sham-HY	GO:0000166	0.0026	nucleotide binding
					GO:0005794	0.0027	Golgi apparatus
thistle1	20	0	5	Sham-HF	No enriched GO terms		

Online Table XII. Fifteen highly connected NTCs with neighboring genes and enriched GO terms.

NTC	module	Ensembl gene	GO ID
chr19:28845241-28862777+	brown	Ace,Col4a1,Igfbp7,Zyx,Tubb5,Pam,Cd9,Islr,Tnfrsf12a,Pfkp,Myh10,Gpx3,Palm2	GO:0006518,GO:0051289,GO:0008360,GO:0001725,GO:0005615,GO:0005200,GO:0007155,GO:0005576,GO:0001666
chr17:86555049-86567124-	greenyellow	Mmp15,Cox5b,Tcap,Tob2,Cox6c,Apba3,Slc41a1,Uqcr,Got2,Unc84b,Vapb,Ndufb7,Art1,Bsg,Cox6a2,Ndufa7,Myl3,Sympk,Gpsm1,Fkrp,Mrpl55,Kcnh2,Ank1	GO:0004129,GO:0005743,GO:0005739,GO:0070469,GO:0060048,GO:0022900,GO:0016529,GO:0008137,GO:0010468,GO:0016020
chr14:55613383-55624058+	greenyellow	Cox5b,Tcap,Tob2,Cox6c,Nfe2l1,Apba3,Slc41a1,Uqcr,Unc84b,Ndufb7,Dbp,Bsg,Myl3,Ndufa7,Cox6a2,Sympk,Gpsm1,Fkrp,Mrpl55,Ank1	GO:0004129,GO:0005743,GO:0005739,GO:0070469,GO:0060048,GO:0022900,GO:0010468,GO:0008137,GO:0016020,GO:0046983
chr13:94318395-94333316-	lightgreen	Trip10,Pcyt1a,Ndufb9,Psap,Cdh4,Mfn2,Gramd4,Slc25a11,Tmem38a,Slc25a3,Nme4,Dusp18	GO:0005743,GO:0005739,GO:0005758,GO:0006626,GO:0008289,GO:0005764,GO:0016020
chr3:53529326-53539395+	purple	Ltbp4,Flnc,Sorbs3,Plxnd1,Atp5f1,Dld,Trim47,Cd74,Qs	GO:0005925,GO:0005737,GO:0003779,GO:0005856,GO:0007266,GO:0045454,GO:0005515

		ox1,Lasp1,D17Wsu104e,Cdc42ep1,Bcar1,BC060632,Prnp,Plekho1,Arhgdia,Tmsb10,Mfge8,Nampt	
chr13:103119664-103127036+	purple	Ltbp4,Pnpla8,Flnc,Sorbs3,Capg,Trim47,Lasp1,D17Wsu104e,Cdc42ep1,Bcar1,Nbl1,Plekho1,Mfge8,Tmsb10,Nampt,Olfml2b	GO:0003779,GO:0005925,GO:0005856,GO:0005737,GO:0005576,GO:0045786
chr2:77148133-77155244-	purple	Flnc,Sorbs3,Tnfrsf1a,Cd74,Bak1,Bcar1,Prnp,Mfge8,Tmsb10	GO:0045121,GO:0006952,GO:0005925,GO:0008283,GO:0005856,GO:0043066,GO:0007155
chr5:114754453-114755202-	purple	Capg,Ptbp2,Trim47,Esrrg,Nbl1,Plekho1,Mcm2,Maff	GO:0045786,GO:0007049,GO:0005634
chr3:63213781-63287912-	red	Efhd2,Fstl3,Serpine1,Spry4,Dynll1,Hsdl2,Csrp1,Spna2,Tln1,Col18a1,Igfbp4,Fth1,Gnb2l1,Col4a2,Serpinh1,Pkd1,Ehd4,Ndufs1,Gpr124,2010111I01Rik,Slc25a25,Xirp1,Csrp3,Rhoc,Scarf2,Ltbp3,Dlg5,Vac14,Vim,Pcolce,Peci,Tubb2a,Nr4a1,Itga5,Cd36,Axl,Actn4,Clu,Cttn,Fam176b	GO:0005916,GO:0005515,GO:0005488,GO:0007044,GO:0032587,GO:0007157,GO:0048812,GO:0005856,GO:0019838,GO:0030863,GO:0005913
chr19:5795590-5802672+	red	Efhd2,Serpine1,Spry4,Dynll1,Hsdl2,Csrp1,Col18a1,Igf1r,Sgk1,Col4a2,Gpr124,Slc25a25,Xirp1,Lrp1,Scarf2,Ltbp3,Vac14,Emilin1,Cd248,Pcolce,Peci,Tubb2a,Nr4a1,Itga5,Cd36,Axl,Pmepa1,Clu,Fam176b	GO:0043277,GO:0002020,GO:0005515,GO:0007157,GO:0005488
chr12:103206306-103214208-	turquoise	Fhl2,Mknk2,Myo18a,Mafg,Atp1a1,	GO:0030641,GO:0022627,GO:0042802,GO:0000287

		Rps27,Phospho1,Rps9,Adipor1,Hspb8,Ptp4a3,Csnk1g2,Adcy5,Atp13a2	
chr10:3131363-3134303-	turquoise	Fhl2,Scamp5,Ddx54,Myo18b,Map3k10,Tns1,Pias4,Gga1,Mafg,Atp1a1,Aldoa,Tenc1,Jph2,Adipor1,Hspb8,Ptp4a3,Atp6v0b,Hand2	GO:0030018,GO:0030641,GO:0042802,GO:0003714,GO:0005925,GO:0009408,GO:0005768
chr10:94934333-94934934-	turquoise	Gse1,Ppp1r12c,Rai12,Wbp1,Nfat5,Ddx3y,Eif2b2,Tmod1,Mnt,Csrp2bp,Pnkd,Arl2bp,Akt1s1,Nme2,Rusc1,Kras,Rbm42,Myo18a,Spns1,Rpsa,Rasgrp2,Ilf3,2310044G17Rik,Heatr2,Znrf1,Rom1,Trafd1,Ccar1,Siva1,Pcif1,Sqstm1,Tcfeb,Man2c1,Rhobtb1,Ap2a2,Ranbp3,Mapt,Klc4,Wnk4,Prkaca,Rere,Irs1,Adipor1,Mapkapk2,Trp53i13,A230051G13Rik,Zfp592,Bat3,Abtb1,Tmem109,Ssscal,Yipf3,Sars,Fhl2,Pip4k2c,Sh2b1,Gabarapl1,Tsc22d4,Map3k10,Tns1,Trappc3,Rnf187,Itga7,Arfgap2,Ldb3,Atp1a1,Neil1,Phf1,Fbxw5,Tollip,Aof2,Rps9,Hspb8,Hdac10,Csnk1g2,Lrrc10,Stk40,Mknk2,Rnf170,Pus1,4930500O05Rik,Erbb2ip,Aqp1,Icam1,Drg2,Saps1,Ppp2r1a,Prkcsh,Ctsl,Ids,Gga1,Ppp1r2,Slc39a14,Ing4,Setd8,150	GO:0005737,GO:0016309,GO:0005080,GO:0005515,GO:0004879,GO:0008113

		0003O22Rik,Kcnq1,D19Wsu162e,Hic1,Xrcc1,Msra,Brap,Pik3c2b,Trak1,Twf2,Sgta,Tenc1,Timp3,Jph2,Dnmt3a,Hmgn2,Taldo1,Map3k3,Vamp7,Hand2,Fbf1,Nif3l1,Dpm3,Mll2,Hdgfrp2,Rab1b,Plekhm2,Phactr4,Sepx1,Rnps1,Dvl3,Gbf1,Pias4,Rnf220,Bnip1,Vcp,Mark4,Pitpnm2,Prdx2,Six5,Taf6l,Pla2g6,Eftud2,Stat6,Cav2,Hmg20b,Rps27,Parp3,Aldoa,Capns1,Mars,Pi4k2a,Exosc4,Pafah2,Mef2d,Epn1,Lzts2,Rpl24,Pip4k2b,Dohh,Fasn,Eif4h,Mink1,Cul9	
chr18:80558615-80562518+	turquoise	CT009486.7-1,Arl2bp,Nme2,Rbm42,Tcfeb,Sqstm1,Bat3,Yipf3,Sars,Fhl2,Ddx54,Pip4k2c,Map3k10,Ldb3,Rps9,Hspb8,Ptp4a3,Atp6v0b,Csnk1g2,Adcy5,Scamp5,Mknk2,Saps1,Gga1,Slc39a14,Rab1b,Sepx1,Myo18b,Pias4,Pitpnm2,Prdx2,Rps27,Aldoa,Epn1,Lzts2	GO:0005737,GO:0030018,GO:0005768,GO:0022627,GO:0005080,GO:0003714,GO:0016055,GO:0005524,GO:0006826,GO:0042802,GO:0009408
chr5:30880396-30882228+	turquoise	Sqstm1,Fhl2,Atp1a1,Hspb8,Ptp4a3,Kcng2,Mafg,Tenc1,Stat6,Phospho1	GO:0030641,GO:0042802,GO:0019901,GO:0042127,GO:0004721,GO:0016791,GO:0005925,GO:0030955

Supplemental References

1. Liao P, Georgakopoulos D, Kovacs A, Zheng M, Lerner D, Pu H, Saffitz J, Chien KR, Xiao R, Kass D, Wang Y. The in vivo role of p38 MAP kinases in cardiac remodeling and restrictive cardiomyopathy. *Proc Natl Acad Sci U S A.* 2001;98:12283-12288.
2. Petrich BG, Molkentin JD, Wang Y. Temporal activation of c-Jun N-terminal kinase in adult transgenic heart via cre-loxP-mediated DNA recombination. *FASEB J.* 2003:02-0438fje.
3. Ruan H, Mitchell S, Vainoriene M, Lou Q, Xie L-H, Ren S, Goldhaber JI, Wang Y. Gi{alpha}1-Mediated Cardiac Electrophysiological Remodeling and Arrhythmia in Hypertrophic Cardiomyopathy. *Circulation.* 2007;116(6):596-605.
4. Rockman HA, Ross RS, Harris AN, Knowlton KU, Steinhelper ME, Field LJ, Ross J, Jr., Chien KR. Segregation of atrial-specific and inducible expression of an atrial natriuretic factor transgene in an in vivo murine model of cardiac hypertrophy. *Proc Natl Acad Sci U S A.* 1991;88(18):8277-8281.
5. Flicek P, Amode MR, Barrell D, Beal K, Brent S, Chen Y, Clapham P, Coates G, Fairley S, Fitzgerald S, Gordon L, Hendrix M, Hourlier T, Johnson N, Kahari A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Larsson P, Longden I, McLaren W, Overduin B, Pritchard B, Riat HS, Rios D, Ritchie GR, Ruffier M, Schuster M, Sobral D, Spudich G, Tang YA, Trevanion S, Vandrovcova J, Vilella AJ, White S, Wilder SP, Zadissa A, Zamora J, Aken BL, Birney E, Cunningham F, Dunham I, Durbin R, Fernandez-Suarez XM, Herrero J, Hubbard TJ, Parker A, Proctor G, Vogel J, Searle SM. Ensembl 2011. *Nucleic Acids Res.* 2011;39(Database issue):D800-806.
6. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.
7. Kent WJ. BLAT--the BLAST-like alignment tool. *Genome Res.* 2002;12(4):656-664.
8. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5(7):621-628.
9. Bullard JH, Purdom E, Hansen KD, Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics.* 2010;11:94.
10. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511-515.
11. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, Weinstock GM, Wilson RK, Gibbs RA, Kent WJ, Miller W, Haussler D. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 2005;15(8):1034-1050.
12. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, Diekhans M, Dreszer TR, Giardine BM, Harte RA, Hillman-Jackson J, Hsu F, Kirkup V, Kuhn RM, Learned K, Li CH, Meyer LR, Pohl A, Raney BJ, Rosenbloom KR, Smith KE, Haussler D, Kent WJ. The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.* 2011;39(Database issue):D876-882.
13. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 2004;11(2-3):377-394.

14. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 2009;19(1):92-105.

15. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell.* 2007;27(1):91-105.

16. Cheng Y, Ji R, Yue J, Yang J, Liu X, Chen H, Dean DB, Zhang C. MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy? *Am J Pathol.* 2007;170(6):1831-1840.

17. Chinchilla A, Lozano E, Daimi H, Esteban FJ, Crist C, Aranega AE, Franco D. MicroRNA profiling during mouse ventricular maturation: a role for miR-27 modulating Mef2c expression. *Cardiovasc Res.* 2011;89(1):98-108.

18. Liu G, Ding M, Chen J, Huang J, Wang H, Jing Q, Shen B. Computational analysis of microRNA function in heart development. *Acta Biochim Biophys Sin (Shanghai).* 2010;42(9):662-670.

19. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. The Pfam protein families database. *Nucleic Acids Res.* 2010;38(Database issue):D211-222.

20. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 2007;35(Web Server issue):W345-349.

21. Vizcaino JA, Cote R, Reisinger F, Foster JM, Mueller M, Rameseder J, Hermjakob H, Martens L. A guide to the Proteomics Identifications Database proteomics data repository. *Proteomics.* 2009;9(18):4276-4283.

22. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.

23. Hulsen T, de Vlieg J, Alkema W. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics.* 2008;9:488.

24. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, Quackenbush J. TM4 microarray software suite. *Methods Enzymol.* 2006;411:134-193.

25. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics.* 2011;27(3):431-432.

26. Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci.* 1998;23(6):198-199.

27. Isken O, Maquat LE. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Dev.* 2007;21(15):1833-1856.

Chapter Three:

Functional Characterization of Novel Exon in PKCα In Heart

Functional Characterization of A Novel Cardiac Specific Splicing Variant in PKCα

Chen Gao[1, 2,3], Gongjian Li[4], Andrea Matlock[2, 3], Jae-Hyung Lee[1, 5] ,Thomas Vondriska[1,2,3], Douglas L. Black[1], Xinshu Xiao[1,5], Susan Steiberg[4], Yibin Wang[1,2,3]

1. The Molecular Biology Institute, UCLA
2. Cardiovascular Research Laboratories, UCLA
3. Division of Molecular Medicine, Departments of Anesthesiology, Physiology and Medicine, David Geffen School of Medicine, UCLA
4. Department of Pharmacology, Columbia University Medical Center
5. Department of Integrative Biology and Physiology, UCLA

Summary:

During our previous RNA-Sequencing study in mouse normal and failing hearts, we have identified 1000+ previously un-annotated novel exons based on different splicing events. Among these, we have identified a novel exon in PKCα, which inserted right before the Turn motif---PKCα-NE (Novel Exon). This novel exon is a cardiac and skeletal muscle specific alternative splicing event, which is at least partially, regulated by RBFox1 based on both in vivo and in vitro analysis. At baseline, the PKCα-NE showed a significant higher phosphorylation level comparing to PKCα-WT. However, this novel isoform has different response towards cardiac hypertrophy stress in vitro, including PMA, Iso and AngII, and has different kinase activity towards classic PKCα downstream targets, including TnI. Based on our mass spececTrometry analysis, we have identified a unique interacting partner for PKCα-NE—eEF1A1. In vitro immune-precipitation analysis demonstrated that PKCα-NE interacts with eEF1A1 and phosphorylates this protein.

Introduction:

The protein kinase C (PKC) family is a critical regulator of cardiac signal transduction. Based on their activation mechanism, the PKC family members are divided into conventional PKC isozymes including PKCα, which respond to Calcium and lipid activation; while the novel and atypical isozymes are Calcium independent but can be activated by lipid(Eric Chruchill 2008, Steinberg 2008). Among the PKC family members, PKCα has been suggested to be the dominant isozyme expressing in mouse and rabbit heart(Ping, Zhang et al. 1997, Pass, Gao et al. 2001), and its expression is dynamically regulated during pathological conditions(Hamplova B 2010). Earlier studies have demonstrated a critical role of PKCα mediating cardiac function and cardiomyocytes contractility. In vitro study using human ventricular cardiomyocytes suggested PKCα translocation from cytosol to the contractile system plays an important role maintaining the contractile force of cardiomyocytes by phosphorylating its downstream target—cardiac Troponin I (TnI)(Molnár, Borbély et al. 2009), in vivo pharmacological inhibition of PKCα using ruboxistaurin has been demonstrated to have antagonizing effect on heart failure post myocardial infarction injury, potentially by regulating cardiac contractility, myocyte cellular contractility, Calcium transients and sarcoplasmic reticulum Calcium load(Hambleton, Hahn et al. 2006, Liu, Chen et al. 2009, Ladage, Tilemann et al. 2011). Further study using transgenic animal models also provided a regulatory mechanism of PKCα in cardiomyocytes hypertrophy via regulating systolic and diastolic function through βAR pathway(Hahn, Marreez et al. 2003).

The regulatory mechanism of PKCα is also well established at post-translational modification level. The newly synthesized PKC has an open conformation that allows the PDK-1 to phosphorylate its priming phosphorylation site, in order to be further activated, the cPKC, including PKCα needs two additional phosphorylation events, one in turn motif and the other in hydrophobic motif. These auto-phosphorylation events are critical for PKCα activity by affecting the enzyme thermal stability, detergent solubility as well as protease/phosphatase susceptibility(Edwards and Newton 1997, Steinberg 2008).

Interestingly, our earlier study based on deep RNA-sequencing on pressure-overload induced mouse failing heart and sham operated normal heart has identified a previously un-annotated novel exon in PKCα(Lee, Gao et al. 2011). Indeed, a significant number of cardiac genes undergo alternative splicing during cardiac development and disease. Including the giant protein Titin, which has been proved to be regulated by a splicing regulator RBM20(Wei Guo 2010, Guo, Schafer et al. 2012, Li, Guo et al. 2012) and SCN5A(Wahbi, Algalarrondo et al. , Murphy, Moon-Grady et al. 2012, Jr 2013). Despite earlier studies focusing on PKCα regulation at total gene expression and phosphorylation level, the alternative splicing regulation of PKCα remains to be explored.

Our RNA-Sequencing study, however, has identified a novel regulatory mechanism of PKCα at alternative splicing level. Based on our deep RNA-Sequencing analysis, we have found a previously un-annotated novel exon of PKCα inserted right in front of the original protein turn motif, and will have a significant impact on the protein structure based on protein structure prediction. In

vitro and in vivo studies have shown the insertion of novel exon in PKCα generated a significant higher level of auto-phosphorylation at turn motif at baseline. Surprisingly, this PKCα-NE (Novel Exon) also has very unique activation profile in cultured cardiomyocytes upon PMA (phorbol 12-myristate 13-acetate), Isoproterenol and Angiotensin II stimulation. Moreover, in contrast to PKCα-WT, the PKCα-NE failed to phosphorylate classic PKCα downsteam target-TnI upon Angiotensin II treatment. In order to determine the binding partner and downstream target of PKCα-NE, we performed immuno-precipitation study followed by mass spectrometry. Interestingly, comparing to PKCα-WT, the PKCα-NE has its unique interacting partners, including key components of protein translation machinery—eEF1A1. We have further demonstrated that eEF1A1 is indeed interacting with PKCα-NE based on immune-precipitation study in cultured cardiomyocytes; and can be potentially phosphorylated by PKCα-NE. Lastly, we investigated the regulatory mechanism of PKCα novel exon alternative splicing. Based on bioinformatics analysis, this highly conserved novel exon also shares conserved flanking cis-regulatory elements across different species. Within the conserved cis-regulatory elements, we have identified two putative RBFox1 binding motifs. Both in vivo and in vitro minigene reporter analysis have provided evidence that RBFox1 indeed regulates the PKCα novel exon splicing directly.

In summary, our deep RNA-sequencing analysis has identified a previously un-annotated cardiac specific splicing event of PKCα. Functional characterization suggested the insertion of this novel exon could have a significant impact on the

enzyme activation responding to hypertrophic stimuli and its downstream target, thus providing a novel regulatory mechanism for this well establish cardiac signal regulator.

Results:

PKCα Novel Exon is A Cardiac And Skeletal Muscle Specific Splicing Event

One advantage of deep RNA-Seq approach is that it allows us to discover previously unknown transcript isoforms. Using the guided transcriptome reconstruction method that we developed in this sequencing project, we have identified a total of 1873 novel exons corresponding to different types of alternative splicing events. Among them, we have chosen novel exon in PKCα as example and carried out further functional characterization. Bioinformatics analysis showed this novel exon is highly conserved across different species (data not shown). Semi-quantitative RT-PCR among different mouse tissues showed this novel exon is a cardiac and skeletal muscle specific alternative splicing event (**Figure 3.1A**). This novel exon inserted right in front of the turn motif (**Figure 3.1B**). This 48bp novel exon will not disrupt the reading frame of the original PKCα, but the insertion of this novel exon will generate different protein structure that will likely interrupt the lipid or Calcium signaling response of PKCα (**Figure 3.1C**).

PKCα-Novel Exon Alternative Splicing Is Regulated By RBFox1

Using bioinformatics analysis, we have analyzed the sequence of PKCα novel exon and the conserved cis-regulatory elements flanking the novel exon. Based on

the sequence information, we have identified two putative RBFox1 binding motifs right next to PKCα novel exon (**Figure 3.2A**). To investigate the alternative splicing regulation of this PKCα novel exon, we used H9C2 cells as in vivo system and carried out gain-of-function analysis. Overexpression of RBFox1 is sufficient to promote the inclusion of PKCα novel exon based on real-time PCR analysis (**Figure** 3.**2B**). To further determine the direct regulatory mechanism of this novel exon, we constructed minigene reporters with different cis-regulatory elements using a previously described fluorescent reporter construct (**Figure 3.2C**). In NIH3T3 cells, we have demonstrated the full-length minigene reporter contains the necessary component for its cardiac splicing specificity (**Figure 3.2D**). Using gain-of-function analysis, we have further demonstrated that overexpression of RBFox1 in NIH3T3 cells could promote inclusion of this novel exon in full-length minigene reporter and the minigene reporter containing at least one RBFox1 binding motif, providing evidence that RBFox1 is indeed a direct trans-activating regulator for the PKCα novel exon alternative splicing (**Figure 3.2E**).

PKCα-NE has Higher Phosphorylation Activity At Baseline

Because the insertion of the novel exon locates right in front of the PKCα turn motif (**Figure 3.1B**), we set to determine whether this novel exon will have any impact on PKCα phosphorylation. Using radiolabeled recombinant PKCα-WT and PKCα–NE, we have identified that the PKCα-NE has a significant higher auto-phosphorylation level comparing to WT-PKCα (**Figure 3.3A**). Using HEK293 cells, we further carried out in vivo gain-of-function analysis showing

overexpression of PKCα-NE has a higher phosphorylation activity on the 638 site comparing to WT-PKCα at baseline (**Figure 3.3B**). This could because the insertion of the novel exon changed the original PKCα structure, making it easier to be phosphorylated at this site; or the insertion of the novel exon generated a novel phosphorylation site that can also be recognized by the p-638 PKCα antibody.

PKCα–NE Failed To Respond To PMA Stimulation in Neonatal Cardiomyocytes

In order to further investigate the functional impact of the novel exon insertion into PKCα, we tested the PKCα-NE response to different hypertrophic stimuli in vitro cultured rat neonatal ventricular cardiomyocytes. Using adenovirus mediated gene delivery, we overexpressed either FLAG-PKCα-WT or FLAG-PKCα-NE in cardiomyocytes, short-term treatment with PMA successfully induced FLAG-PKCα-WT translocation from cytosol to plasma membrane and peri-nuclei region (**Figure 3.4A**), however, FLAG-PKCα-NE failed to respond to PMA stimulation and remained majorly in the cytosol fraction. Previous study has suggested PMA treatment would also cause PKCα degradation, our result showed that FLAG-PKCα-NE does not degrade even after long time PMA treatment, opposing to the FLAG-PKCα-WT (**Figure 3.4B**).

PKCα-NE has Different Activation Profile Upon ISO and AngII Stimulation and Does Not Phosphorylate Classic PKCα Substrate

To determine the kinase activity and substrate preference of PKCα-NE, we also tested PKCα-NE kinase activity towards classic PKCα substrate both in vitro and in vivo. We chose radiolabeled CREB peptide as it has been previously suggested to be a downstream target for PKCα. Upon PMA stimulation, in vitro purified PKCα-WT successfully phosphorylated radiolabeled CREB peptide, however, PKCα-NE failed to phosphorylate this classic PKCα substrate under both baseline and PMA stimulation (**Figure 3.5A**).

Using cultured cardiomyocytes, we also tested whether PKCα-NE could phosphorylate endogenous PKCα substrate. Upon isoproterenol stimulation, PKCα-WT is activated through phosphorylation on S657 site while the PKCα-NE is activated through phosphorylation on T638 site (**Figure 3.5B**), suggesting a different activation profile between WT and NE PKCα upon isoproterenol treatment. Lastly, we tested PKCα-NE response to Angiotensin II in cardiomyocytes. According to **Figure 3.5C**, Angiotensin II treatment also activates PKCα-WT via phosphorylation on S657 site, while activating PKCα-NE via phosphorylation on T638 site. Interestingly, PKCα-WT successfully phosphorylated endogenous downstream target—TnI, but this target is not phosphorylated by PKCα-NE (**Figure 3.5C**).

Both PKCα-WT and PKCα-NE Induced Hypertrophy Response in NRVM

Previous study has suggested a critical role of PKCα mediating hypertrophy response in both cultured cardiomyocytes and in intact heart. To test the functional impact of novel exon insertion in PKCα on cardiomyocytes

hypertrophy, we overexpressed individual PKCα isoforms in cultured cardiomyocytes and investigated the cardiomyocytes hypertrophy response. Comparing to mock infected cardiomyocytes, overexpressing PKCα-WT in NRVM induced hypertrophy response based on increased cell size. Surprisingly, overexpressing PKCα-NE stimulated cardiomyocytes hypertrophy response and had an even more significant increase of cell size comparing to PKCα-WT (**Figure 3.6A**). We further tested the hypertrophy marker gene expression level among mock, PKCα-WT and PKCα-NE infected cardiomyocytes. Interestingly, both PKCα-WT and PKCα-NE induced expression of ANF and βMHC to the same level (**Figure** 3.**6B**). Taken together, these data provided evidence that PKCα-NE has different activation profile upon hypertrophic stress stimulation, and potentially has different downstream targets and interacting partners.

Mass Spectrometry Analysis Identified Novel Interacting Partners for PKCα-NE

In order to identify the novel interacting partners with PKCα-NE, we performed immune-precipitation followed by Mass spectrometry. FLAG-PKCα-WT and FLAG-PKCα-NE were individually overexpressed in neonatal cardiomyocytes and precipitated with FLAG antibody (Figure **3.7A**). Interestingly, the PKCα-WT and PKCα-NE have different interacting partners according to Mass Spectrometry results. While PKCα-WT majorly interacted with previously identified chaperon proteins including Hsp70 and Hsp90, PKCα-NE has been demonstrated to

interacte with additional partners, including key regulators for protein synthesis -- eEF1A1(**Figure3.7B**).

eEF1A1 is A Novel PKCα-NE Interacting Partner and is Potentially Phosphorylated by PKCα-NE

In order to evaluate whether eEF1A1 is indeed interacting with PKCα-NE, we performed immune-precipitation in NRVM by overexpressing FLAG-PKCα-NE and compared with FLAG-PKCα-WT. We confirmed our Mass spectrometry result by showing eEF1A1 is interacting with only PKCα-NE but not PKCα-WT in NRVM (**Figure 3.8A**). To further demonstrate this interaction is specific, we generated a mutant construct by replacing the novel exon insertion into alanine mutant. Interestingly, in HEK293 cells, PKCα-WT did not interact with eEF1A1, further, PKCα-NE-Ala mutant failed to interact with eEF1A1 as well, suggesting that it is the unique sequence of the novel exon insertion that is required for the interaction. To further demonstrate that eEF1A1 is a potential downstream target, we probed the blot with phosphor-PKCα substrate antibody. **Figure 3.8B** demonstrated that, PKCα-NE interacted with eEF1A1 specifically, and this interaction also leads to eEF1A1 phosphorylation.

Discussion:

In this study, we have identified a cardiac and skeletal muscle specific splicing variant for PKCα, the insertion of this novel exon locates right in front of the turn motif and will have a potential functional impact on the original protein structure. In vitro biochemistry characterization and in vivo cell study has demonstrated this novel splicing variant of PKCα ---PKCα-NE has a significant higher auto-

phosphorylation level at baseline. PKCα-NE also has a different activation profile in cultured cardiomyocytes upon PMA, AngII and ISO treatment, potentially due to its unique interacting partners and downstream substrates as suggested by mass spectrometry study.

Moreover, we have also provided evidence based on both in vivo and in vitro study that this alternative splicing event in PKCα is at least partially regulated by RBFox1.

A New Layer of Cardiac Transcriptome Complexity

During the past decade, with deep RNA-Sequencing technology, we have obtained a more comprehensive understanding of mammalian transcriptome at the level of single base resolution. Comparing to previous microarray study, more transcriptome components have been established to be functionally important, including alternative splicing, microRNA, lncRNA, piRNA, novel transcript clusters(Lappalainen, Sammeth et al. 2013, Severino, Oliveira et al. 2013, Sun, You et al. 2013, Weikard, Hadlich et al. 2013).

In order to provide a more comprehensive understanding of the cardiac transcriptome, as well as the dynamics of cardiac transcriptome under normal and pathological conditions, we performed an earlier study using RNA-sequencing technology to determine the transcriptome complexity in pressure-overload induced mouse hypertrophy and failing hearts, and compared with sham-operated control hearts(Lee, Gao et al. 2011). From this study, we have confirmed previous study showing there are a significant number of genes differentially expressed between diseased and normal hearts. We have also identified more than one

thousand genes showing differential alternative splicing between failing and normal hearts. Interestingly, taking advantage of de novo annotation feature of RNA-sequencing(Adamidi 2011), we have further identified a significant number of lncRNA and novel transcript clusters with dynamic expression profile tightly associated with cardiac disease. Lastly, we have identified a large amount of previously un-annotated novel splicing variants. However, majority of these novel splicing variants are expressed at relatively low level. Thus, the functional impact of these novel splicing variants remains to be explored.

In this study, we used the novel splicing variant in PKCα as example and carried out detailed biochemistry and cell study. Our data provided strong evidence that this novel splicing variant is dynamically expressed during cardiac development and pathological conditions (data-not-shown). It is also a tightly regulated alternative splicing event carried out by cardiac and skeletal muscle enriched splicing regulator—RBFox1 (**Figure 3.3**). Both biochemistry and cell study suggested this novel splicing variant also has different maturation and activation profile comparing to original PKCα post hypertrophy stimulation.

Thus, our study has provided evidence that there are still a significant number of previously uncharacterized functional components in cardiac transcriptome, including novel splicing variants. These novel splicing variants will lead to novel protein products that might have distinct functional impact during cardiac normal and pathological conditions. Future study on these under-explored novel splicing variants, including using large scale proteomics study to identify the absolute expression level of these novel splicing variants at protein level, would provide us

a more comprehensive understanding of the cardiac transcriptome complexity and the dynamics of cardiac transcriptome during development and disease progression.

A Novel Regulatory Mechanism of PKCα

Among the PKC kinase family members, PKCα is the dominant isozyme expressed in heart(Steinberg 2008). Previous studies have majorly focusing on regulating this kinase at total expression level and phosphorylation level(Bayer AL 2003, Koide, Tamura et al. 2003, Steinberg 2008, Hamplova B 2010), including an up-regulation of this kinase during hypertrophy and failing heart, as well as a series of phosphorylation events in both turn motif and hydrophobic motif that ultimately activate PKCα. Upon activation, the translocation of PKCα is also tightly regulated in order for the kinase to phosphorylate its downstream substrate, including cardiac troponin I(Rybin, Xu et al. 1999, Itoh, Ding et al. 2005, Chakraborti, Roy et al. 2013).

Our study, on the other hand, has provided a novel regulatory mechanism of PKCα at alternative splicing level. The insertion of this 48bp novel exon will not disrupt the protein reading frame, but rather changes the protein structure close to the turn motif phosphorylation site (**Figure** 3.**1C**). In vitro biochemistry and in vivo cardiomyocytes study both suggested this novel splicing variant of PKCα has distinct activation and maturation profile comparing to WT PKCα, based on its auto-phosphorylation activity as well as its phosphorylation and translocation profile in response to hypertrophy stimuli (**Figure 3.4-5**).

In summary, our results showed, in addition to different isozymes in PKC family having different roles and activation profile during cardiomyocytes hypertrophy, the splicing variants in PKCα can also have distinct activation profile during maturation. This splicing variant is cardiac specific (**Figure 3.1A**) and is tightly regulated during cardiac development and pathological remodeling. Further study on this novel exon will provide novel insights on PKCα regulation and its functional impact during cardiac pathological remodeling.

In order to further dissect the functional difference between PKCα-NE and PKCα-WT, future study can focus on analyzing crystal structure of PKCα-NE and PKCα-WT and compare their different response towards Calcium signaling and lipid signaling. Although biochemistry study and cell study have both pointed out a significant higher auto-phosphorylation level of PKCα-NE at baseline, this could due to the insertion of the novel exon augmented the turn motif phosphorylation, or it can be contributed by the potential phosphorylation site in novel exon itself. Additional study generating phosphorylation dead mutant in PKCα-NE novel exon will provide a better understanding of the kinase phosphorylation profile.

A New Hypertrophy Pathway in Cardiomyocytes

According to previous study, PKCα is the dominant isozyme among PKC kinase family in mediating cardiomyocytes contractility and hypertrophic response(Hahn, Marreez et al. 2003, Liu, Chen et al. 2009). This kinase is the dominant kinase expressed in adult cardiomyocytes and its expression is dynamically regulated under different physiological and pathological

conditions(Rybin, Xu et al. 1999, Hamplova B 2010). The PKCα also stands in the center of multiple signaling pathways, including MAPK kinase, ROS and Calcinuerin pathways(De Windt, Lim et al. 2000, Itoh, Ding et al. 2005, Chakraborti, Roy et al. 2013). Thus, a lot of pharmacological inhibitors have been developed for PKCα, inhibiting both kinase activity and its downstream signaling pathways(Hambleton, Hahn et al. 2006, Eric Chruchill 2008, Ladage, Tilemann et al. 2011, Haarberg, Li et al. 2013).

This study, on the other hand, characterized in detail a novel splicing variant of PKCα specifically expressed in cardiac and skeletal muscle (**Figure 3.1A**). In order to provide a molecular mechanism underlying the differential activation profile between PKCα-NE and PKCα-WT upon hypertrophic stimuli, we performed mass spectrometry study following cardiomyocytes immune-precipitation (**Figure** 3.**7A**). Surprisingly, the PKCα-NE indeed has unique interacting partners, including key components in protein translation machinery—eEF1A1 (**Figure** 3.**7B**). Further in vivo study using both cardiomyocytes and HEK293 cells demonstrated that, not only eEF1A1 indeed interacts with PKCα-NE, it can also be potentially phosphorylated by eEF1A1 (**Figure 3.8**). These evidence suggested, in addition to the well characterized original PKCα-WT, we have established a novel PKCα splicing variant that have distinct interacting partner and downstream target. Considering the hypertrophy response induced by PKCα-NE (**Figure 3.6**) and different activation profile of this PKCα-NE responding to ISO and AngII stimulation, PKCα-NE can also have significant value as a therapeutic target in heart failure and cardiac hypertrophy. Apparently,

this PKCα-NE has its unique downstream pathway involving protein synthesis machinery, targeting this novel pathway could provide alternative therapy for cardiac hypertrophy.

Future study in this part involves fine mapping of the interaction site between PKCα-NE and eEF1A1. Both cardiomyocytes and HEK293 cells study supported the interaction between PKCα-NE and eEF1A1. It is not clear, however, whether it is the novel exon provides interacting surface with eEF1A1, or the insertion of the novel exon changes the original protein structure, therefore generates a novel interaction site within the original protein. Additional study will focus on generating alanine mutant insertion into this novel exon and mapping the real interaction site between PKCα-NE and eEF1A1. Another question is, whether eEF1A1 is indeed phosphorylated by PKCα-NE specifically, and whether this phosphorylation event is required for the interaction. Based on literature search, we indeed found a perfect phosphorylation site in eEF1A1 that can be potentially phosphorylated by PKCα-NE—Ser 52 site. Further study will include mutating this particular site in eEF1A1 and carry out further immune-precipitation assay. This will provide us a more detailed molecular mechanism on how this PKCα-NE interacts and regulates eEF1A1. The mapping of the interaction site on PKCα-NE would also provide insights on inhibitor design.

Experimental Procedure and Methods:

Reagents:

PMA (Phorbal 12-myristate 13-acetate) was purchased from Sigma-Aldrich (P8139). Cardiomyocytes was treated with PMA for 30min or 8hr as previously

described(Kirchhefer, Heinick et al. 2013). Angiotensin II was purchased from Sigma-Aldrich (A2900) and cardiomyocytes were treated as previously described(Herrera, Silva et al. 2010). Isoproterenol was purchased from Sigma-Aldrich (I6504) and cardiomyocytes were treated as previously described(Chakraborti, Roy et al. 2013).

Molecular Cloning and Minigene Reporter:

The pShuttle-CMV vector was linearized using HindIII and KpnI. PKCa-WT and PKCa-NE was amplified using primers with the same restriction site (see **Table 1**). Following PCR, the product was purified and digested with HindIII and KpnI before ligated into the linearized pShuttle-CMV vector.

Minigene reporter with fluorescent was a kind gift from Dr. DL Black in UCLA. The minigene reporter was constructed as previously described(Zheng, Damoiseaux et al. 2013) using primers included in **Table 1**. NIH3T3 cells were transfected with individual minigene reporter alone or in combination with RBFox1 expression vector. 48hr post transfection, cells were visualized under fluorescent microscopy to determine GFP/RFP expression level.

Adenovirus:

Adenovirus for FLAG-PKCα-WT and FLAG-PKCα-NE were prepared using AdEasy Adenoviral Vector System (Stratagene) according to manufacturer's protocol. The generation of pShuttle vectors and virus preparation has been described earlier(Lu, Ren et al. 2007).

Cell Culture, Transfection and Western Blot:

HEK293 cells and NIH3T3 cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen) with 10% FBS at 37C and 5% CO2 according to standard ATCC protocol. Cells were transfected with pShuttle-CMV-FLAG-PKCα-WT construct or pShuttle-CMV-FLAG-PKCα-NE construct using Lipofectamine 2000 reagent (Invitrogen) according to manufacturer's protocol. 48hrs post transfection, cells were harvested for protein extraction. Western Blot was performed using Novex NuPAGE Gel Electrophoresis Systems (Invitrogen) according to manufacturer's protocol. 40ug protein was used per sample. The auto-phosphorylation level of PKCα is determined by anti-PKCα (Phospho T638) and anti-PKCα (phosphor S657+Y658) (Abcam), or anti-PKCα (Cell Signaling). β-Actin (Santa Cruz Biotechnology) is used as the internal loading control.

NRVM (Neonatal Rat Ventricular Myocytes) was isolated from p1 neonatal rat and plated on Gelatin (Sigma-Aldrich) coated 6-well dish. Cells were maintained in Dulbecco's modified Eagle's medium (Invitrogen) with 1% ITS solution (BD Biosciences) 24hr before adenovirus infection. Cells were collected 48hr post infection, and protein was isolated for Western Blot analysis using Phospho-Troponin I (Cardiac Ser 23/24) (Cell Signaling) and Troponin I Antibody (Cell Signaling).

Real-time PCR:

RNA was isolated from cells using 1ml Trizol Reagent (Invitrogen) according to manufacturer's protocol. cDNA was synthesized from 0.5ug RNA using random Hexamer and SuperScript II RT (Invitrogen) according to manufacturer's

protocol. 0.5ul cDNA was used per 20ul real-time PCR reaction using SYBR Green supermix (Bio-Rad) with primers specific for PKCα-WT or PKCα-NE.

Immuno-fluorescence:

NRVM was plated on glass-bottom dish (in vitro scientific) and infected with FLAG-PKCα-WT or FLAG-PKCα-NE. 24hr post infection, cells were treated with PMA. 30min post PMA treatment, NRVM was fixed in 4% Formaldehyde at room temperature for 15min, followed by 3 times wash with ice-cold PBS. Cells was blocked with blocking buffer (1*PBS/5% normal goat serum/0.3% Triton X-100) at room temperature for 1hr before incubating with FLAG antibody (SigmaAldrech) at 4 degree overnight. After washing cells with ice cold PBS 3 times, cells were further incubated with secondary Alexa Fluor 488 goat anti mouse antibody (Invitrogen) at room temperature for 1hr. Cells were further washed with ice-cold PBS before confocal microscopy analysis.

Immuno-precipitation:

NRVM was infected with FLAG-PKCα-WT or FLAG-PKCα-NE individually. 48hr post infection, cells were collected with protein lysis buffer (20mM Tris pH8.0; 137mM NaCl, 0.5%NP-40; 2mM EDTA; PMSF;protease inhibitor cocktail tablet). 1ug protein was pre-cleared with normal serum on ice for 1hr before adding 50ul Anti-FLAG M2 affinity gel (Sigma-Aldrich A2220) and incubate at 4 degree overnight. The affinity gel was washed with washing buffer (10mM Tris pH7.4, 1mM EDTA; 1mM EGTA;150mM NaCl, 0.5%NP40;PMSF; protease inhibitor cocktail) for a total of 6 times before elution with FLAG peptide (Sigma-Aldrich F4799) according to manufacturer's protocol. After

elution, samples were denatured using NuPAGE LDS Sample Buffer (Invitrogen). The SDS-PAGE was visualized by incubating with Oriole Stain (Bio-Rad) at room temperature for 2hrs; or directly proceeds with Western Blot followed by detection using eEF1A1 antibody (Abcam).

Mass spectrometry:

Mass spectrometry was performed as previously described(Lu, Sun et al. 2009). Generally, proteins were digested with trypsin and analyzed by LC/MS/MS on a Thermo LTQ Orbitra mass spectrometer with an Eksigent NanoLC pump. The peptides were loaded onto a C18 reverse-[hase and 20% water in CAN. Peptides were eluted from the column at a flow rate of 220nl/min, using a linear gradient from 5% B to 50% B over 90 minutes, then to 95% B over 5 minutes, and finally keeping constant 95% B for 5 minutes. Spectra were acquired in data-dependent mode with the Orbitrap used for mass spectrometry scans and LTQ for tandem mass spectrometry. Peptides were identified by searching the spectra against the rat International Protein Index using the SEQUEST algorithm integrated into the BioWorks software package.

Statistics:

2-tailed Student's t test was performed to determine the significance of difference between RBFox1 overexpressing H9C2 and control H9C2 cells to determine PKCα alternative splicing profile. $P<0.05$ is considered significant.

Table 3.1: List of Primers

Primer Name	Sequence
Prkcα KpnI F	ATTAGGTACCATGGCTGACGTTTACCCGGCCAAC
Prkca FLAG HindIII R	TTATAAGCTTTCACTTATCGTCGTCATCCTTGTAATCTAC TGCACTTTGCAAGATT
mPrkca RT-F	CCAAGCGGCTGGGCTGCGGG
mPrkca NE RT-R	CTGAAGCCAGTGCATTTTGGT
mPrkca RT-R	GTCAAAGTTTTCTGCTCCTTT
rPrkca RT-F	CGTTCGATGGCGAAGACGAA
rPrkca NER	CCACTGAAGCCAGTGCATTTT
rPrkca WTR	CTTTGCCGCACACTTTGGGC
Prkca EcoRI F1	ATTAGAATTCATGTGACATGCATGTGTAAA
Prkca BamHI R1	TTAAGGATCCGGCCAGTGGCAGACACAGGCT
Prkca EcoRI F2	ATTAGAATTCCATACAGCCGTTCCTGATTC
Prkca BamHI R2	TTAAGGATCCTCTGCCCTTTCCATGAAGGT
Prkca BamHI R3	TTAAGGATCCAGGGAAGGTCTTGATGCACGA
Prkca BamHI	TTAAGGATCCGATGAGGGCAAGCAGCTGCAC

R4	
Prkca BamHI R5	TTAAGGATCCCGGTTAATGGTGAAAAGTGAGG
Prkca BamHI R6	TTAAGGATCCGCTTAGTCACTCAAGGTTCTGCT

Figure Legend:

Figure 1: PKCα-Novel Exon is a cardiac specific alternative splicing event

A) Different tissues from adult mouse were collected and RNA was extracted. Semi-quantitative RT-PCR was performed to determine PKCα novel exon existence among different tissues. GAPDH is used as internal control.

B) Schematic view of novel exon insertion in PKCα

C) Protein structure prediction of PKCα-WT and PKCα-NE. Predicted by SWISS-MODEL.

Figure 2: PKCα-NE alternative splicing is regulated by RBFox1

A) Gene structure of PKCα-NE with flanking cis-regualtory elements generated by UCSC genome browser. RBFox1 and ASF/SF1 binding motif is indicated.

B) In vivo PKCα-NE splicing analysis. H9C2 cells were infected with Ade-FLAG-RBFox1 and compared with mock infected cells. Real-time PCR was performed to determine the ratio between PKCα-WT vs PKCα-NE. n=3 each sample, *,$p<0.05$

C) Schematic view of different versions of PKCα-NE minigene reporter

D) PKCα-NE minigene reporter alternative splicing in NIH3T3 cells. NIH3T3 cells were transfected with individual PKCα-NE minigene reporter. 48hr post transfection, cells were visualized under fluorescent microscope.

E) In vitro PKCα-NE alternative splicing regulation. NIH3T3 cells were transfected with individual PKCα-NE minigene reporter alone or in

combination with FLAG-RBFox1 expression vector as indicated. 48hr post transfection, cells were visualized under fluorescent microscope.

Figure 3: PKCα-NE has higher auto-phosphorylation level at baseline

A) FLAG-PKCα-WT and FLAG-PKCα-NE were overexpressed in HEK293 cells and purified using FLAG antibody immune-precipitation. Purified protein was radio-labeled and visualized under phosphor-imager.

B) PKCα-WT and PKCα-NE were overexpressed in HEK293 cells. 48hr post transfection, cells were harvested and Western Blot was performed to determine the phosphorylation level of PKCα-WT and PKCα-NE using phosphor-PKCα antibodies. β-Actin is used as loading control. Quantification represents ratio between p-638 PKCα vs total PKCα.

Figure 4: PKCα-NE has different response to PMA treatment in neonatal cardiomyocytes

A) PKCα-NE translocation upon PMA treatment. NRVM was infected with FLAG-PKCα-WT and FLAG-PKCα-NE adenovirus individually. Cells were treated with PMA for 30min and compared with control NRVM. Immuno-fluorescent was performed using FLAG antibody and translocation was visualized under confocal microscopy.

B) PKCα-NE degradation upon PMA treatment. NRVM was infected with FLAG-PKCα-WT and FLAG-PKCα-NE adenovirus individually. 24hr post infection, cells were treated with PMA for 8hr and compared with control NRVM. Western Blot was performed to determine total protein level of PKCα-WT and PKCα-NE.

Figure 5: PKCα-NE has different substrate and activation profile upon AngII and ISO treatment.

A) PKCα-NE fails to phosphorylate classic PKCα substrate in vitro. FLAG-PKCα-WT and FLAG-PKCα-NE was expressed in HEK293 cells individually and purified by FLAG antibody immune-precipitation. The purified protein was incubated with 35S labeled CREB peptide with PMA treatment and compared with control protein lysates. CREB peptide was visualized using phosphor-imager to determine PKCα-NE kinase activity.

B) PKCα-NE activation profile upon ISO treatment. NRVM was infected with FLAG-PKCα-WT and compared with FLAG-PKCα-NE and mock-infected cells. 24hr post infection, cells were treated with ISO for 8hr. Western Blot was performed to analyze activation profile of PKCα-WT and PKCα-NE using phosphor-PKCα antibody. Actin was used as loading control.

C) PKCα-NE activation profile upon AngII treatment. NRVM was infected with FLAG-PKCα-WT and FLAG-PKCα-NE adenovirus. 24hr post infection, cells were treated with AngII for 30min and compared with control cells. PKCα-NE activation was determined using phosphor-PKCα antibodies. PKCα–NE kinase activity was determined using p-TnI antibody.

Figure 6: PKCα-WT and PKCα-NE both induced cardiomyocytes hypertrophy

A) NRVM was infected with FLAG-PKCα-WT or FLAG-PKCα-NE individually. 48hr post transfection, cells were visualized under bight field microscope to analyze cell size.

B) FLAG-PKCα-WT and FLAG-PKCα-NE were overexpressed in NRVM. 48hr post infection, cells were harvested and real-time PCR was performed to determine the hypertrophy response using ANF and βMHC markers. n=3 each sample, *, $p<0.05$.

Figure 7: Identification of PKCα-NE interaction partner

A) Mass spectrometry analysis of PKCα-WT and PKCα-NE interacting partners**.** FLAG-PKCα-WT and FLAG-PKCα-NE were overexpressed in NRVM via adenovirus. 48hr post infection, cells were harvested and immune-precipitation was performed using FLAG antibody. SDS-PAGE was visualized using Oriole Stain. Protein samples were processed as indicated.

B) Protein ID identified to be interacting with PKCα-WT and PKCα-NE according to mass spectrometry.

Figure 8: eEF1A1 interacts with PKCα-NE

A) eEF1A1 interacts with PKCα-NE in NRVM. NRVM was infected with FLAG-PKCα-WT and FLAG-PKCα-NE. 48hr post infection, cells were harvested and immune-precipitation was performed using FLAG antibody. Western Bot was performed to determine eEF1A1 interaction with PKCα-NE.

B) eEF1A1 is potentially phosphorylated by PKCα-NE. HEK293 cells were transfected with FLAG-PKCα-WT and FLAG-PKCα-NE. 48hr post transfection, cells were collected and immune-precipitation was performed using FLAG antibody. Western Blot was used to determine the phosphorylation of eEF1A1 using phosphor-PKCα-substrate antibody.

Acknowledgements:

The authors would like to thank Dr. Douglas L Black (UCLA) for kindly providing minigene reporter for the minigene reporter analysis. The author would also like to thank Dr. Thomas Vondriska for his discussion and technique support for the mass spectrometry study. The authors also appreciate the discussion and experimental suggestions provided by Dr. Susan Steinberg (Columbia University).

C. G is supported by Eli & Edythe Broad center of Regenerative Medicine &Stem Cell Research Center Training Grant.

References:

Adamidi, C. (2011). "De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics." Genome Res. **21**: 1193-1200.

Bayer AL, H. M., Patel N, Porter M, Engman S, Samarel AM (2003). "Alterations in protein kinase C isoenzyme expression and autophosphorylation during the progression of pressure-overload induced left ventricular hypertrophy." Mol Cell Biochem **243**(2): 145-152.

Chakraborti, S., S. Roy, A. Chowdhury, A. Mandal and T. Chakraborti (2013). "Role of PKCα−p38MAPK−Giα axis in peroxynitrite-mediated inhibition of β-adrenergic response in pulmonary artery smooth muscle cells." Cellular Signalling **25**(2): 512-526.

De Windt, L. J., H. W. Lim, S. Haq, T. Force and J. D. Molkentin (2000). "Calcineurin Promotes Protein Kinase C and c-Jun NH2-terminal Kinase Activation in the Heart: CROSS-TALK BETWEEN CARDIAC HYPERTROPHIC SIGNALING PATHWAYS." Journal of Biological Chemistry **275**(18): 13571-13579.

Edwards, A. S. and A. C. Newton (1997). "Phosphorylation at Conserved Carboxyl-terminal Hydrophobic Motif Regulates the Catalytic and Regulatory Domains of Protein Kinase C." Journal of Biological Chemistry **272**(29): 18382-18390.

Eric Chruchill, G. B., Alice Vallentin, Tomoyoshi Koyanagi, and Daria Mochly-Rosen (2008). "PKC Isozymes in Chronic Cardiac Disease: Possible Therapeutic Targets." Annual Reviews Pharmacology and Toxicology(48): 569-599.

Guo, W., S. Schafer, M. L. Greaser, M. H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A. M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. A. Hacker, K. W. Saupe, G. W. Dec, P. T. Ellinor, C. A. MacRae, B. Spallek, R. Fischer, A. Perrot, C. Ozcelik, K. Saar, N. Hubner and M. Gotthardt (2012). "RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing." Nat Med **18**(5): 766-773.

Haarberg, K. M. K., J. Li, J. Heinrichs, D. Wang, C. Liu, C. C. Bronk, K. Kaosaard, A. M. Owyang, S. Holland, E. Masuda, K. Tso, B. R. Blazar, C. Anasetti, A. A. Beg and X.-Z. Yu (2013). "Pharmacologic inhibition of PKCα and PKCθ prevents GVHD while preserving GVL activity in mice." Blood **122**(14): 2500-2511.

Hahn, H. S., Y. Marreez, A. Odley, A. Sterbling, M. G. Yussman, K. C. Hilty, I. Bodi, S. B. Liggett, A. Schwartz and G. W. Dorn (2003). "Protein Kinase Cα Negatively Regulates Systolic and Diastolic Function in Pathological Hypertrophy." Circulation Research **93**(11): 1111-1119.

Hambleton, M., H. Hahn, S. T. Pleger, M. C. Kuhn, R. Klevitsky, A. N. Carr, T. F. Kimball, T. E. Hewett, G. W. Dorn, W. J. Koch and J. D. Molkentin (2006). "Pharmacological- and Gene Therapy-Based Inhibition of Protein Kinase Cα/β Enhances Cardiac Contractility and Attenuates Heart Failure." Circulation **114**(6): 574-582.

Hamplova B, N. F., Kolar F, Novakova O (2010). "Transient upregulation of protein kinase C in pressure-overloaded neonatal rat myocardium
." Physiological Research **59**(1): 25-33.

Herrera, M., G. B. Silva and J. L. Garvin (2010). "Angiotensin II Stimulates Thick Ascending Limb Superoxide Production via Protein Kinase Cα-dependent NADPH Oxidase Activation." Journal of Biological Chemistry **285**(28): 21323-21328.

Itoh, S., B. Ding, C. P. Bains, N. Wang, Y. Takeishi, T. Jalili, G. L. King, R. A. Walsh, C. Yan and J.-i. Abe (2005). "Role of p90 Ribosomal S6 Kinase (p90RSK) in Reactive Oxygen Species and Protein Kinase C β (PKC-β)-mediated Cardiac Troponin I Phosphorylation." Journal of Biological Chemistry **280**(25): 24135-24142.

Jr, G. G. a. S. C. D. (2013). "SCN5A splicing variants and the possibility of predicting heart failure-associated arrhythmia." Expert Reveiw of Cardiovascular Therapy **11**(2): 117-119.

Kirchhefer, U., A. Heinick, S. König, T. Kristensen, F. U. Müller, M. D. Seidl and P. Boknik (2013). "Protein phosphatase 2A is regulated by PKCα-dependent phosphorylation of its targeting subunit B56α at Ser41." Journal of Biological Chemistry.

Koide, Y., K. Tamura, A. Suzuki, K. Kitamura, K. Yokoyama, T. Hashimoto, N. Hirawa, M. Kihara, Ohno, Shigeo and S. Umemura (2003). "<B>Differential Induction of Protein Kinase C Isoforms at the Cardiac Hypertrophy Stage and

Congestive Heart Failure Stage in Dahl Salt-Sensitive Rats</B>." Hypertension Research **26**(5): 421-426.

Ladage, D., L. Tilemann, K. Ishikawa, R. N. Correll, Y. Kawase, S. R. Houser, J. D. Molkentin and R. J. Hajjar (2011). "Inhibition of PKCα/β With Ruboxistaurin Antagonizes Heart Failure in Pigs After Myocardial Infarction Injury." Circulation Research **109**(12): 1396-1400.

Lappalainen, T., M. Sammeth, M. R. Friedlander, P. A. C. t Hoen, J. Monlong, M. A. Rivas, M. Gonzalez-Porta, N. Kurbatova, T. Griebel, P. G. Ferreira, M. Barann, T. Wieland, L. Greger, M. van Iterson, J. Almlof, P. Ribeca, I. Pulyakhina, D. Esser, T. Giger, A. Tikhonov, M. Sultan, G. Bertier, D. G. MacArthur, M. Lek, E. Lizano, H. P. J. Buermans, I. Padioleau, T. Schwarzmayr, O. Karlberg, H. Ongen, H. Kilpinen, S. Beltran, M. Gut, K. Kahlem, V. Amstislavskiy, O. Stegle, M. Pirinen, S. B. Montgomery, P. Donnelly, M. I. McCarthy, P. Flicek, T. M. Strom, C. The Geuvadis, H. Lehrach, S. Schreiber, R. Sudbrak, A. Carracedo, S. E. Antonarakis, R. Hasler, A.-C. Syvanen, G.-J. van Ommen, A. Brazma, T. Meitinger, P. Rosenstiel, R. Guigo, I. G. Gut, X. Estivill and E. T. Dermitzakis (2013). "Transcriptome and genome sequencing uncovers functional variation in humans." Nature **501**(7468): 506-511.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research **109**(12): 1332-1341.

Li, S., W. Guo, B. M. Schmitt and M. L. Greaser (2012). "Comprehensive analysis of titin protein isoform and alternative splicing in normal and mutant rats." Journal of Cellular Biochemistry **113**(4): 1265-1273.

Liu, Q., X. Chen, S. M. MacDonnell, E. G. Kranias, J. N. Lorenz, M. Leitges, S. R. Houser and J. D. Molkentin (2009). "Protein Kinase Cα, but Not PKCβ or PKCγ, Regulates Contractility and Heart Failure Susceptibility: Implications for Ruboxistaurin as a Novel Therapeutic Approach." Circulation Research **105**(2): 194-200.

Lu, G., S. Ren, P. Korge, J. Choi, Y. Dong, J. Weiss, C. Koehler, J.-n. Chen and Y. Wang (2007). "A novel mitochondrial matrix serine/threonine protein phosphatase regulates the mitochondria permeability transition pore and is essential for cellular survival and development." Genes & Development **21**(7): 784-796.

Lu, G., H. Sun, P. She, J.-Y. Youn, S. Warburton, P. Ping, T. M. Vondriska, H. Cai, C. J. Lynch and Y. Wang (2009). "Protein phosphatase 2Cm is a critical regulator of branched-chain amino acid catabolism in mice and cultured cells." The Journal of Clinical Investigation **119**(6): 1678-1687.

Molnár, A., A. Borbély, D. Czuriga, S. M. Ivetta, S. Szilágyi, Z. Hertelendi, E. T. Pásztor, Á. Balogh, Z. Galajda, T. Szerafin, K. Jaquet, Z. Papp, I. Édes and A. Tóth (2009). "Protein Kinase C Contributes to the Maintenance of Contractile Force in Human Ventricular Cardiomyocytes." Journal of Biological Chemistry **284**(2): 1031-1039.

Murphy, L. L., A. J. Moon-Grady, B. F. Cuneo, R. T. Wakai, S. Yu, J. D. Kunic, D. W. Benson and A. L. George Jr (2012). "Developmentally regulated SCN5A splice variant potentiates dysfunction of a novel mutation associated with severe fetal arrhythmia." Heart Rhythm **9**(4): 590-597.

Pass, J. M., J. Gao, W. K. Jones, W. B. Wead, X. Wu, J. Zhang, C. P. Baines, R. Bolli, Y.-T. Zheng, I. G. Joshua and P. Ping (2001). "Enhanced PKCβII translocation and PKCβII-RACK1 interactions in PKCε-induced heart failure: a role for RACK1." American Journal of Physiology - Heart and Circulatory Physiology **281**(6): H2500-H2510.

Ping, P., J. Zhang, Y. Qiu, X.-L. Tang, S. Manchikalapudi, X. Cao and R. Bolli (1997). "Ischemic Preconditioning Induces Selective Translocation of Protein Kinase C Isoforms ε and η in the Heart of Conscious Rabbits Without Subcellular Redistribution of Total Protein Kinase C Activity." Circulation Research **81**(3): 404-414.

Rybin, V. O., X. Xu and S. F. Steinberg (1999). "Activated Protein Kinase C Isoforms Target to Cardiomyocyte Caveolae: Stimulation of Local Protein Phosphorylation." Circulation Research **84**(9): 980-988.

Severino, P., L. Oliveira, N. Torres, F. Andreghetto, M. de Fatima Guarizo Klingbeil, R. Moyses, V. Wunsch-Filho, F. Nunes, M. Mathor, A. Paschoal and A. Durham (2013). "High-throughput sequencing of small RNA transcriptomes reveals critical biological features targeted by microRNAs in cell models used for squamous cell cancer research." BMC Genomics **14**(1): 735.

Steinberg, S. F. (2008). "Structural Basis of Protein Kinase C Isoform Function." Physiological Reviews **88**(4): 1341-1378.

Sun, W., X. You, A. Gogol-Doring, H. He, Y. Kise, M. Sohn, T. Chen, A. Klebes, D. Schmucker and W. Chen (2013). "Ultra-deep profiling of alternatively spliced Drosophila Dscam isoforms by circularization-assisted multi-segment sequencing." EMBO J **32**(14): 2029-2038.

Wahbi, K., V. Algalarrondo, H. M. Bécane, V. Fressart, C. Beldjord, K. Azibi, A. Lazarus, N. Berber, H. Radvanyi-Hoffman, T. Stojkovic, A. Béhin, P. Laforêt, B. Eymard, S. Hatem and D. Duboc "Brugada syndrome and abnormal splicing of SCN5A in myotonic dystrophy type 1." Archives of Cardiovascular Diseases(0).

Wei Guo, S. J. B., Karla Esbona and Marion L. Greaser (2010). "Titin Diversity-Alternative Splicing Gone Wild." J Biomed Biotechnol.

Weikard, R., F. Hadlich and C. Kuehn (2013). "Identification of novel transcripts and noncoding RNAs in bovine skin by deep next generation sequencing." BMC Genomics **14**(1): 789.

Zheng, S., R. Damoiseaux, L. Chen and D. L. Black (2013). "A broadly applicable high-throughput screening strategy identifies new regulators of Dlg4 (Psd-95) alternative splicing." Genome Research **23**(6): 998-1007.

Figure 3: Graphic Abstract

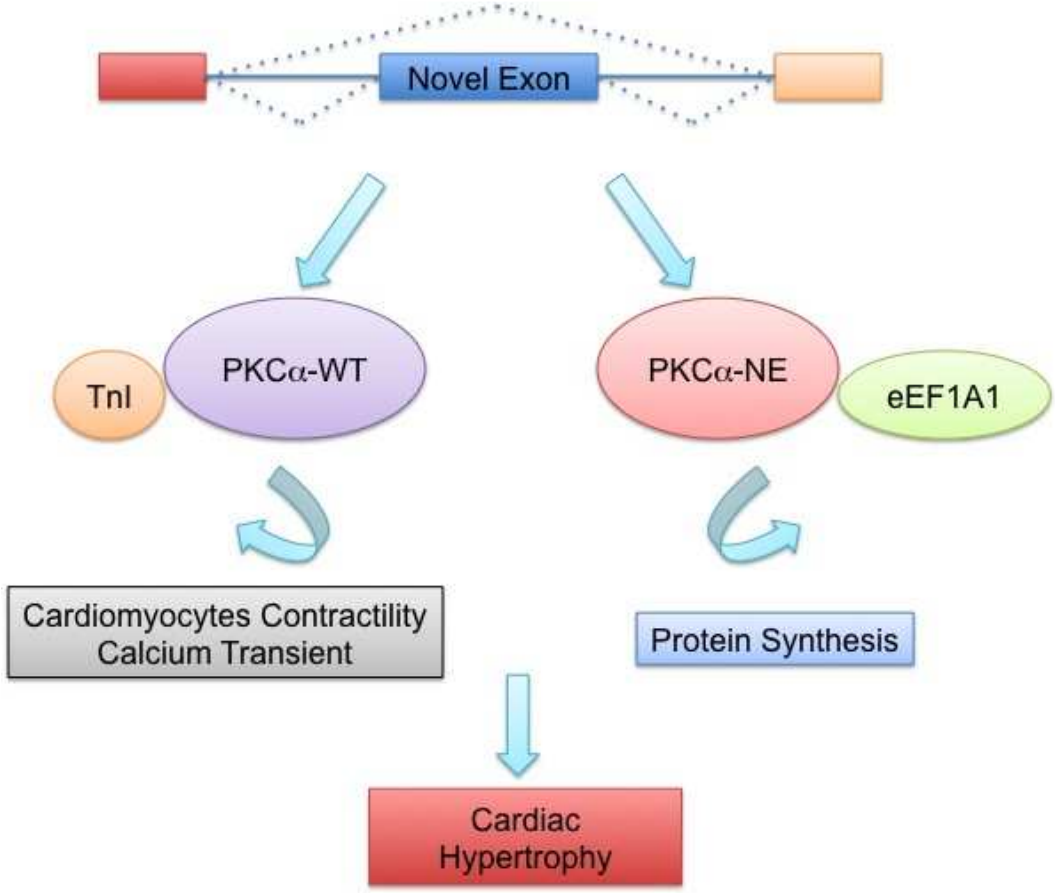
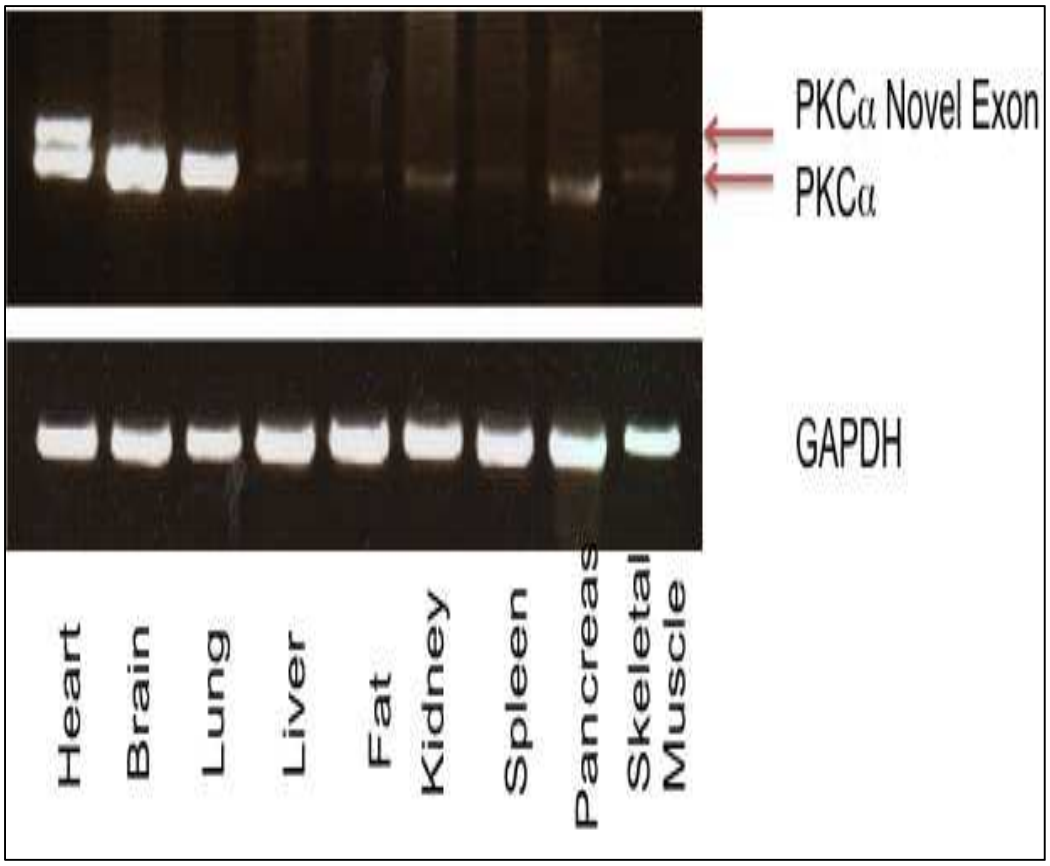


Figure 3.1: PKCα-NE is a cardiac and skeletal muscle specific alternative splicing event



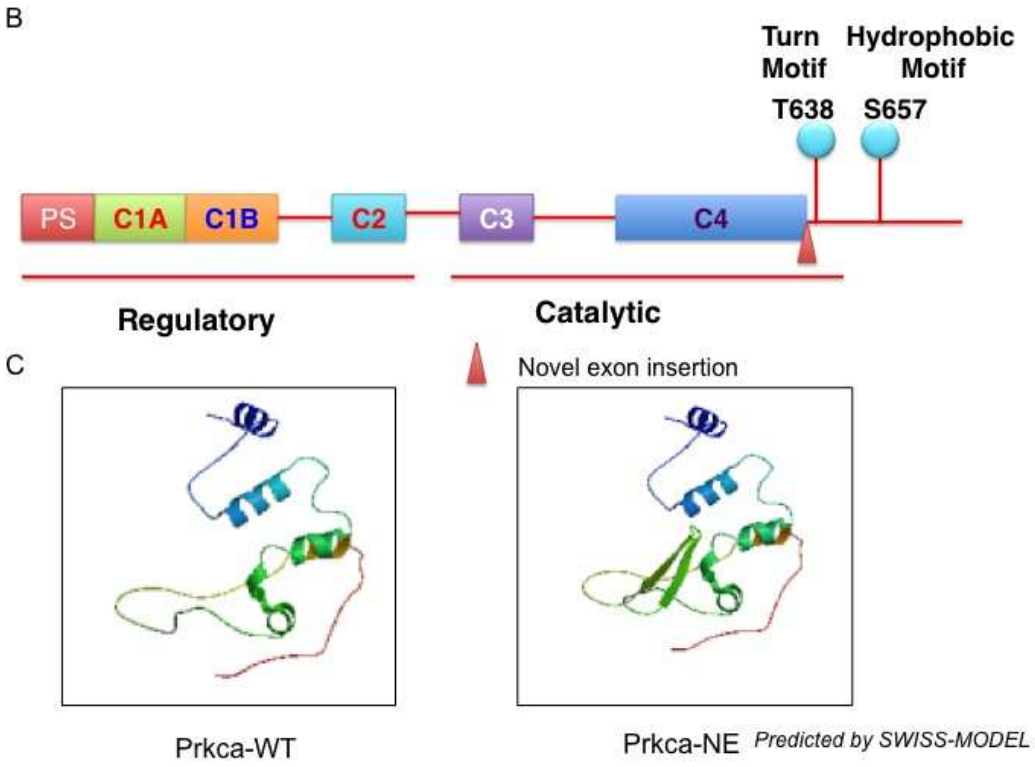
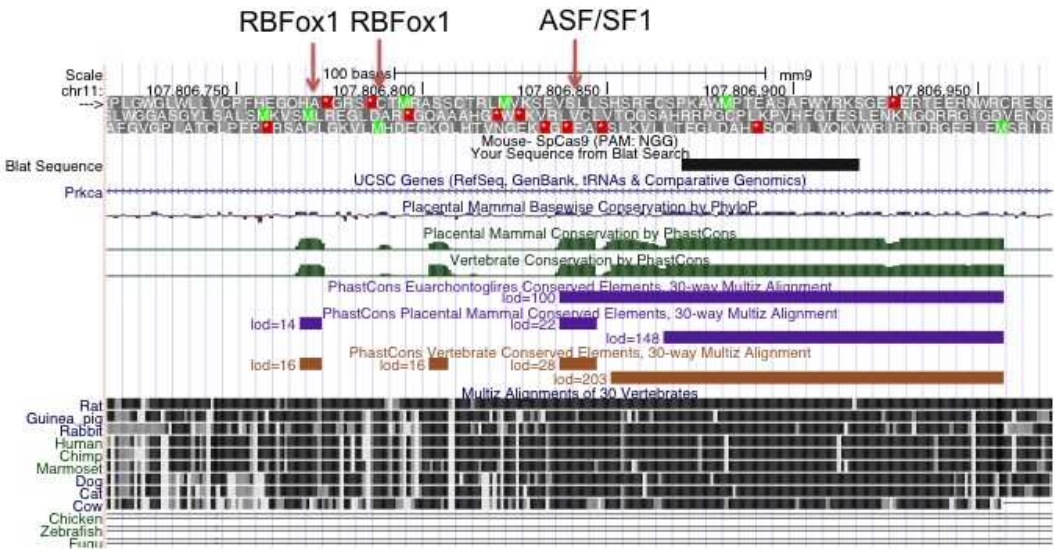

B
Turn Motif
Hydrophobic Motif
T638
S657
PS
C1A
C1B
C2
C3
C4
Regulatory
Catalytic
C
Novel exon insertion
Prkca-WT
Prkca-NE
Predicted by SWISS-MODEL

Figure 3.2 PKCα-NE alternative splicing is regulated by RBFox1



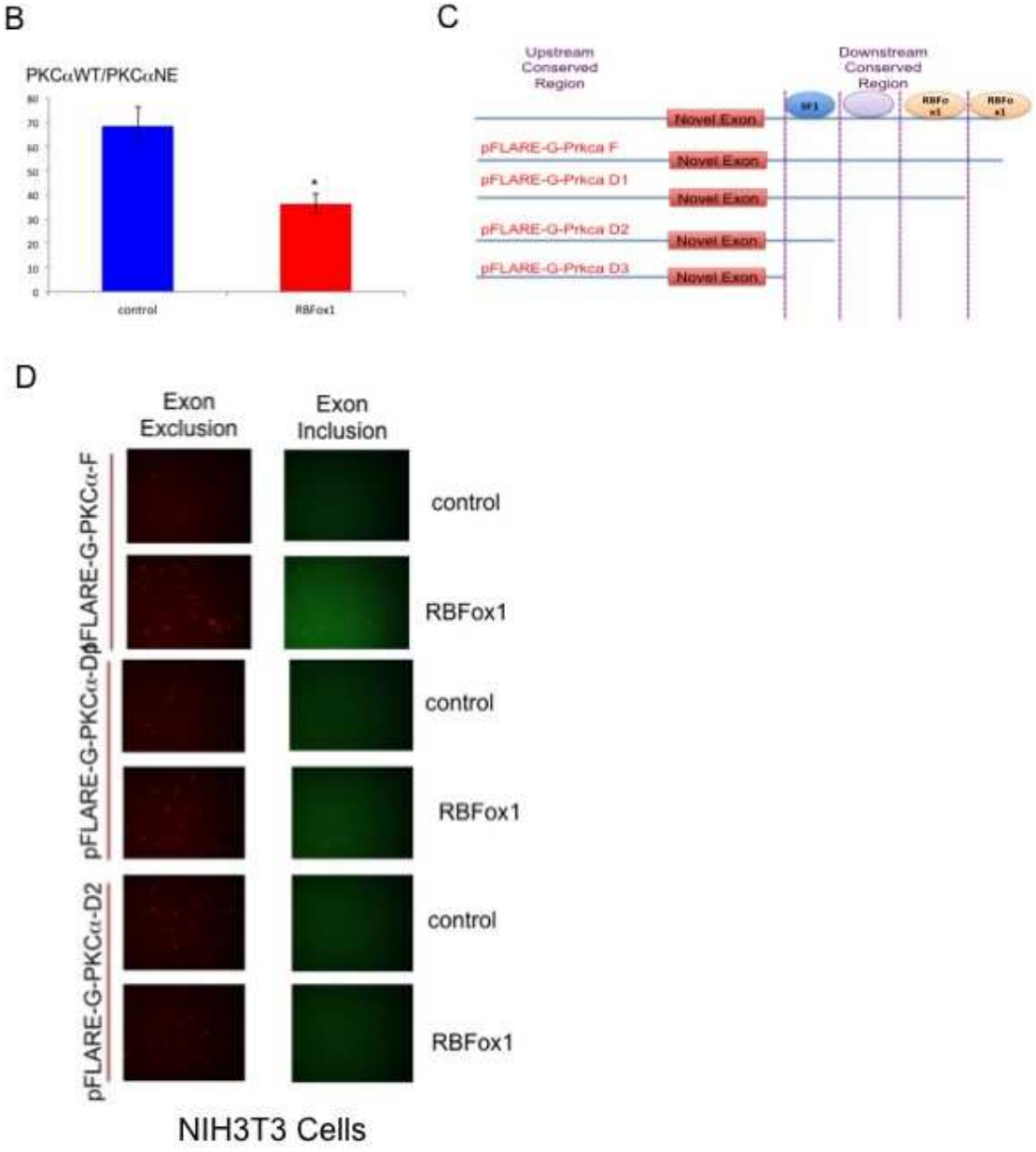

B
PKCαWT/PKCαNE
control
RBFox1
C
Upstream Conserved Region
Downstream Conserved Region
Novel Exon
pFLARE-G-Prkca F
pFLARE-G-Prkca D1
pFLARE-G-Prkca D2
pFLARE-G-Prkca D3
D
Exon Exclusion
Exon Inclusion
pFLARE-G-PKCα-F
pFLARE-G-PKCα-D1
pFLARE-G-PKCα-D2
control
RBFox1
NIH3T3 Cells

Figure 3.3 PKCα-NE has higher auto-phosphorylation level at baseline

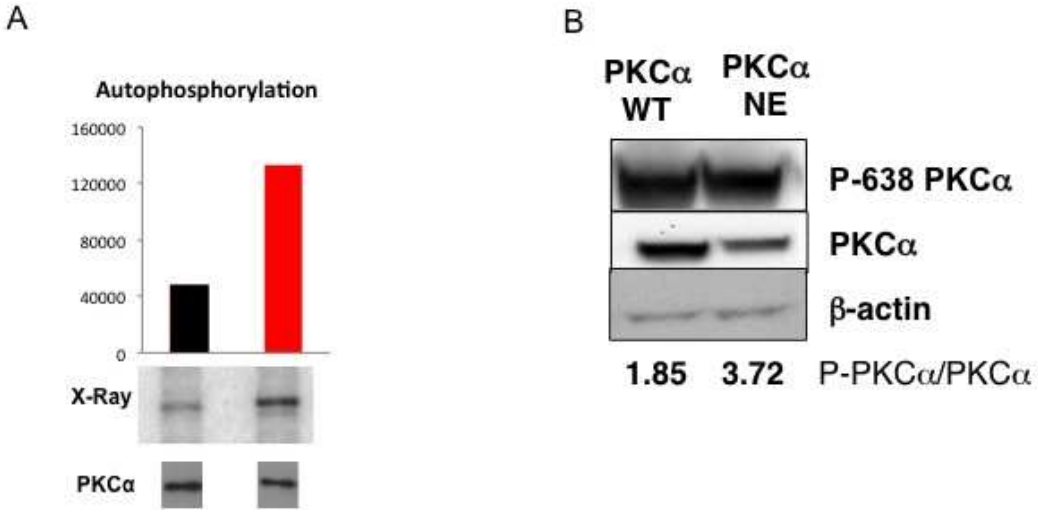


Figure 3.4 PKCα-NE has different activation profile responding to PMA stimulation

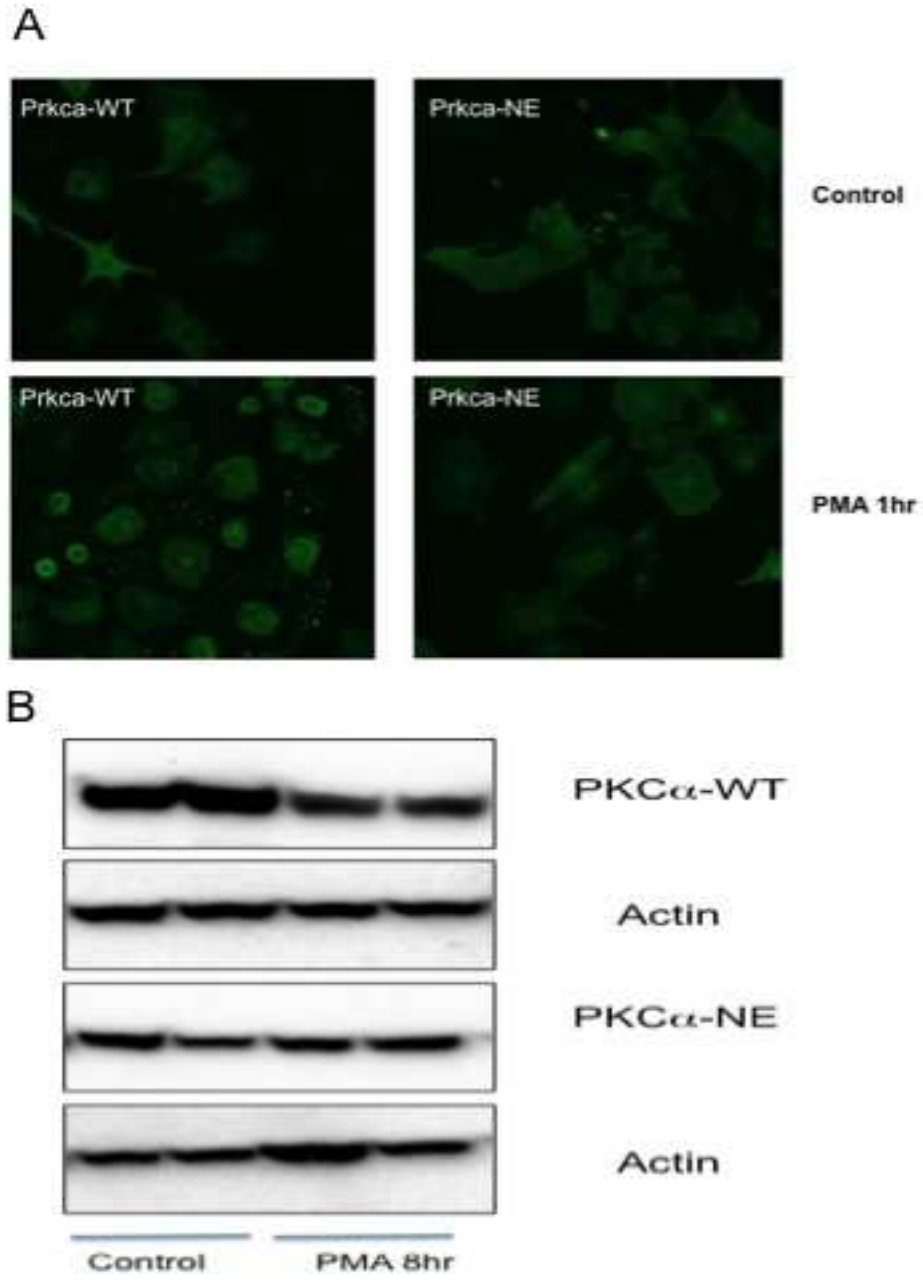


Figure 3.5 PKCα-NE has different activation in response to ISO and AngII treatment in NRVM

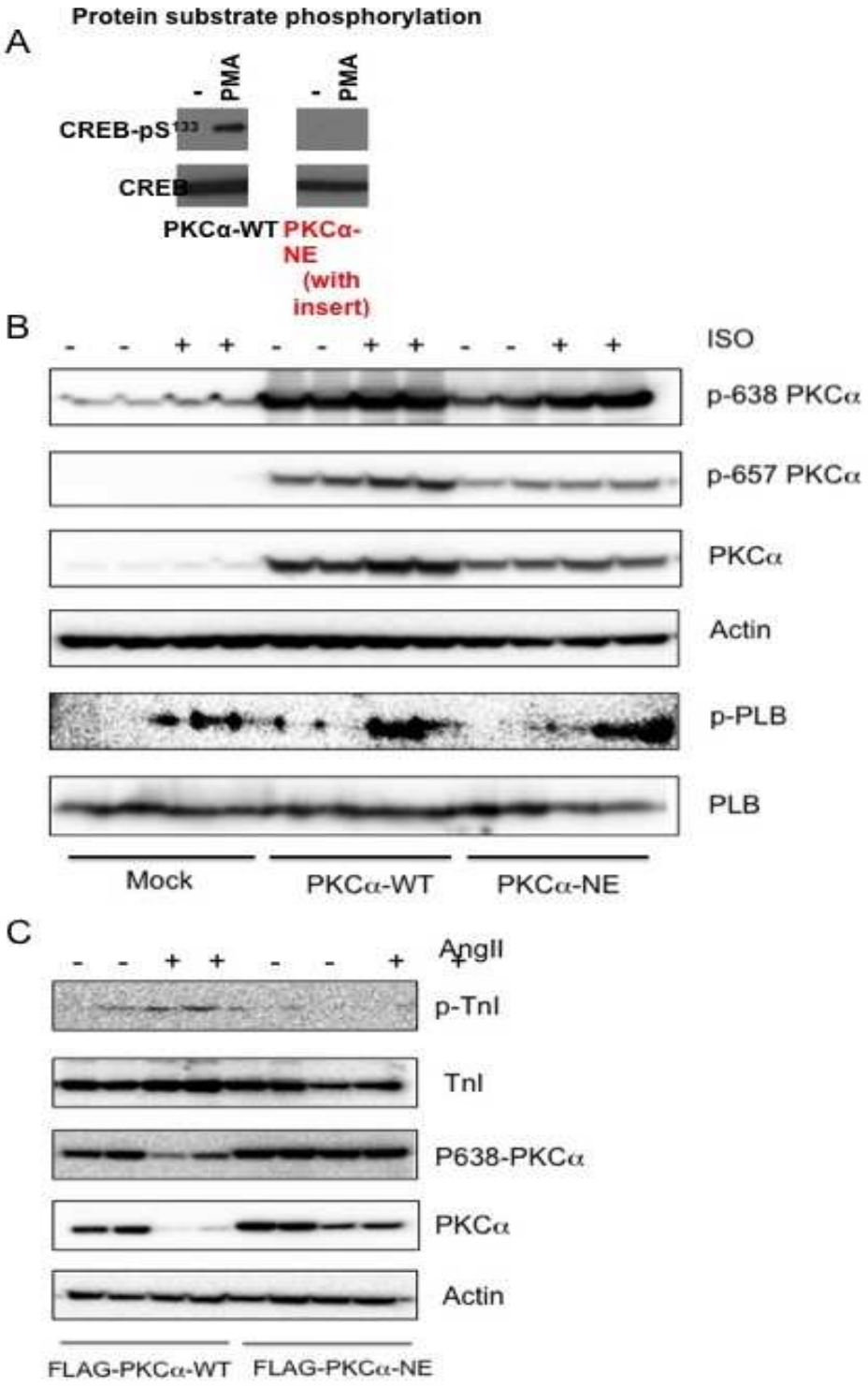


Figure 3.6 PKCα-WT and PKCα-NE both induced hypertrophy response in cardiomyocytes

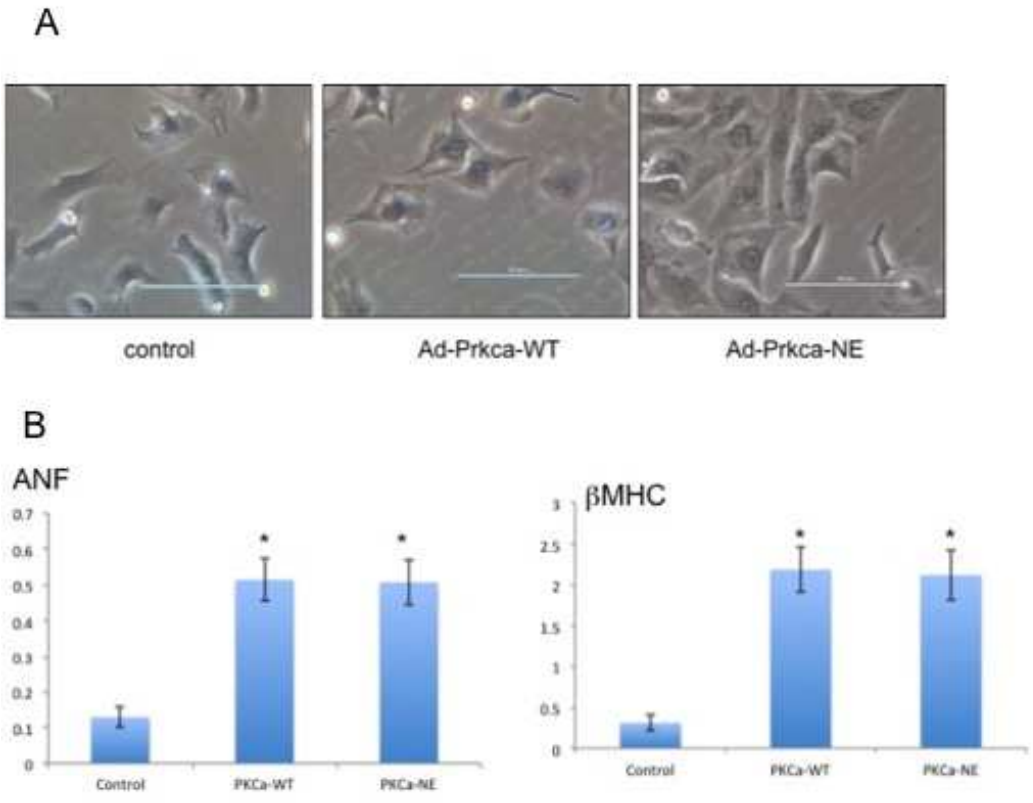
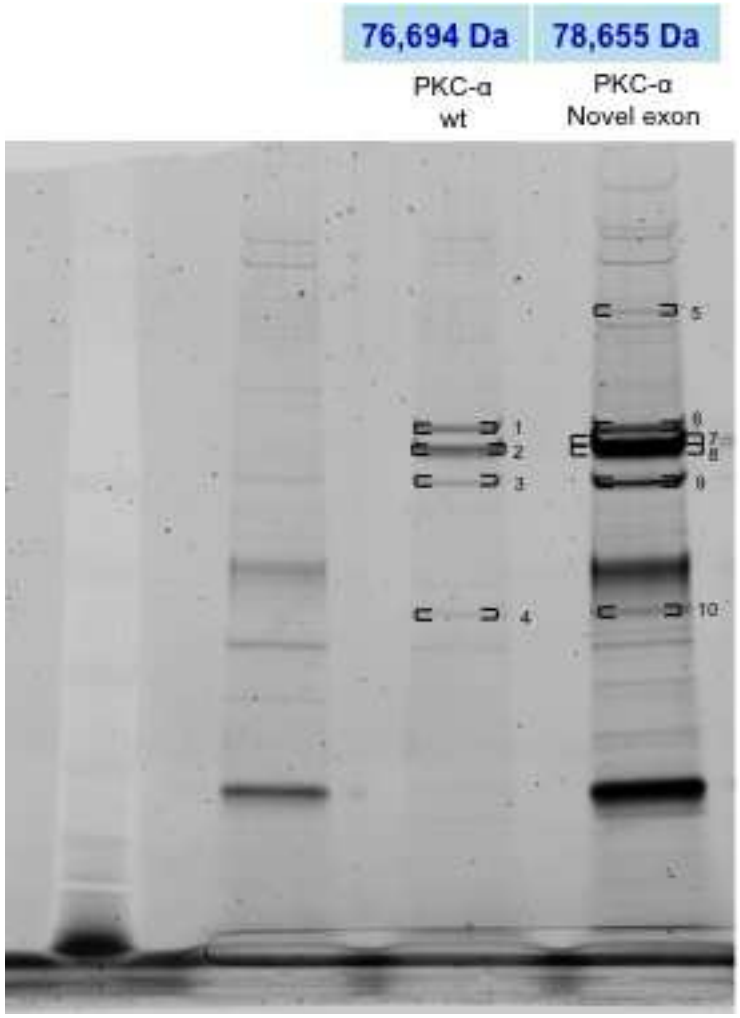
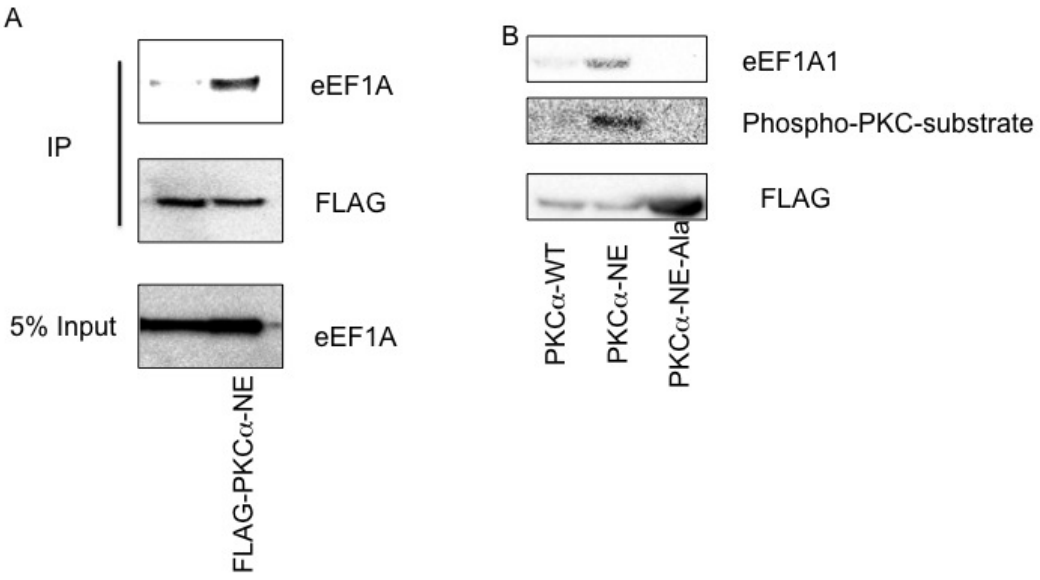


Figure 3.7 Mass spectrometry identified novel interaction partners for PKCα-NE



Gel Band	Protein	Gel band	Protein
1	Isoform 2 of Heat shock protein HSP 90-alpha	5	Eukaryotic translation initiation factor 5B
	Heat shock protein HSP 90-beta		Rho-associated protein kinase 2
2	Protein kinase C alpha type		Polyubiquitin-C
3	Protein arginine N-methyltransferase 5	6	Isoform 2 of Heat shock protein HSP 90-alpha
	CALL5_HUMAN Calmodulin-like protein 5		Heat shock protein HSP 90-beta
	Stress-70 protein, mitochondrial		Inactive serine/threonine-protein kinase PLK5
	Heat shock 70 kDa protein 1A/1B	7	Protein kinase C alpha type
	Heat shock 70 kDa protein 6	8	BAG family molecular chaperone regulator 3
	78 kDa glucose-regulated protein		Protein kinase C alpha
	Heat shock cognate 71 kDa protein	9	ATP-dependent RNA helicase DDX3X
4	Isoform 2 of 26S proteasome non-ATPase regulatory subunit 11		Protein arginine N-methyltransferase 5
	Hsp90 co-chaperone Cdc37		Heat shock 70 kDa protein 1A/1B
	Isoform 2 of Protein disulfide-isomerase A6		Stress-70 protein, mitochondria
	Elongation factor Tu, mitochondrial		Heat shock cognate 71 kDa protein
	Elongation factor 1-alpha 1		Heat shock 70 kDa protein 6
		10	Elongation factor 1-alpha 1
			Elongation factor Tu, mitochondrial
			Heterogeneous nuclear ribonucleoprotein F

Figure 3.8 PKCα-NE interacts and potentially phosphorylates eEF1A1



Chapter Four

Cardiac Genes Undergo Extensive Alternative Splicing

Introduction:

In the eukaryotic transcriptome, both the numbers of genes and different RNA species produced by each gene contribute to the overall complexity. These RNA species are generated by the utilization of different transcriptional initiation or termination sites, or more commonly, from different messenger RNA (mRNA) splicing events. Among the 30,000+ genes in human genome, it is estimated that more than 95% of them can generate more than one gene product via alternative RNA splicing. The protein products generated from different RNA splicing variants can have different intracellular localization, activities, or tissue-distribution. Therefore, alternative RNA splicing is an important molecular process that contributes to the overall complexity of the genome and the functional specificity and diversity among different cell types. In this review, we have discussed current efforts to unravel the full complexity of the cardiac transcriptome using a deep-sequencing approach, and highlighted the potential of this technology to uncover the global impact of RNA splicing on the transcriptome during development and diseases of the heart.

Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Perspective:

Global impact of RNA splicing on transcriptome remodeling in the heart*

Chen GAO, Yibin WANG†‡
(*Molecular Biology Institute, University of California, Los Angeles, California 90095, USA*)
†E-mail: yibinwang@mednet.ucla.edu
Received June 20, 2012; Revision accepted July 13, 2012
Crosschecked July 19, 2012

doi: 10.1631/jzus.B1201006
Document code: A **CLC number:** R4; Q78

In the eukaryotic transcriptome, both the numbers of genes and different RNA species produced by each gene contribute to the overall complexity. These RNA species are generated by the utilization of different transcriptional initiation or termination sites, or more commonly, from different messenger RNA (mRNA) splicing events. Among the 30000+ genes in human genome, it is estimated that more than 95% of them can generate more than one gene product via alternative RNA splicing. The protein products generated from different RNA splicing variants can have different intracellular localization, activity, or tissue-distribution. Therefore, alternative RNA splicing is an important molecular process that contributes to the overall complexity of the genome and the functional specificity and diversity among different cell types. In this review, we will discuss current efforts to unravel the full complexity of the cardiac transcriptome using a deep-sequencing approach, and highlight the potential of this technology to uncover the global impact of RNA splicing on the transcriptome during development and diseases of the heart.

Key words: Alternative RNA splicing, Transcriptome, Gene regulation, Heart, RNA-seq

‡ Corresponding author
* Project supported partially by Broad Stem Cell Research Center (BSCRC) Pre-doctoral Fellowship in UCLA, and the National Institutes of Health, USA

1 Alternative RNA splicing

The complexity of the eukaryotic transcriptome was first fully revealed at the genome scale with single-base resolution by a powerful deep RNA-sequencing (RNA-seq) technology, using next generation sequencers and newly developed bioinformatic tools (Bland *et al*., 2010; Hallegger *et al*., 2010). It is estimated that transcripts from ~95% of multi-exon genes undergo alternative splicing and that there are ~100000 intermediate to high abundance alternative splicing events in major human tissues (Pan *et al*., 2008). RNA splicing is a ubiquitous post-transcriptional process in all eukaryotes. It involves removing intronic sequences from pre-messenger RNA (pre-mRNA) and linking exons to generate mature mRNA for translation (Chen and Manley, 2009). RNA splicing for constitutively spliced exons is carried out by a defined molecular machinery involving cis-acting regulatory sequences (splice sites) located at exon-intron boundaries, as well as trans-acting factors as part of the spliceosome (de la Grange *et al*., 2010). However, in many cases, the splice sites are altered, leading to different exon sizes in the final transcripts. Alternatively, certain exons can be differentially included or excluded in the final transcripts due to exon skipping. These non-constitutive RNA splicing activities are collectively called alternative RNA splicing. Through these different kinds of alternative pre-mRNA processing, individual eukaryotic genes can produce multiple mRNA and protein isoforms that may have related, distinct or even opposing functions (Wang *et al*., 2008; Buljan *et al*., 2012). Therefore, alternative RNA splicing is an important molecular step that contributes to the total complexity of the transcriptome and proteome.

2 Regulation of alternative RNA splicing

Alternative RNA splicing is a highly regulated process mediated by cis-regulatory enhancers and silencers in pre-mRNA and trans-acting splicing factors, including heterogeneous nuclear ribonucleoprotein (hnRNP) and serine-arginine rich proteins (SR proteins). The molecular nature of these regulatory elements is yet to be fully uncovered and understood. Tissue specific alternative splicing is usually regulated by a combination of tissue-specific and ubiquitously expressed splicing factors (Pan *et al*., 2008; Sultan *et al*., 2008; Chen and Manley, 2009; Bland *et al*., 2010) and has been demonstrated to play an important role in regulation of tissue-specific protein interaction networks (Buljan *et al*., 2012). In addition, mis-regulated alternative RNA splicing events have a significant role in human diseases, cell cycle, and cell death (Hallegger *et al*., 2010; Gang *et al*., 2011; Honda *et al*., 2012; Raghavachari, 2012; Yae *et al*., 2012). For example, a specific isoform of pyruvate kinase resulting from hnRNP-mediated mRNA alternative splicing is required for tumor cell proliferation (David *et al*., 2010).

3 Alternative RNA splicing in cardiac diseases

It is well established that alternative splicing of mRNA is tightly associated with the development of heart failure. Structural proteins, such as cardiac troponin T, or important signaling molecules, such as Ca^{2+}/calmodulin-dependent protein kinase (CaM kinase), are subjected to alternative splicing in heart diseases (Ramchatesingh *et al*., 1995; Ding *et al*., 2004; Xu *et al*., 2005). Moreover, depletion of critical splicing regulators, including SC35 and RBM20, has been found to cause dilated cardiomyopathy in mouse and rat (Ding *et al*., 2004; Guo *et al*., 2012; Linke and Bucker, 2012; Refaat *et al*., 2012). In addition to classic SR and hnRNP proteins, CUG-BP1 and ETR-like factors (CELF)/Bruno-like family of RNA binding proteins and muscleblind-like proteins (MBNL proteins) have also been found to regulate both cardiac development and function (Warf and Berglund, 2007; Kalsotra *et al*., 2010; Koshelev *et al*., 2010; Dasgupta and Ladd, 2012). Therefore, alternative RNA splicing is essential for normal cardiac function and mis-regulated RNA splicing may have an important role in the pathogenesis of heart failure. Yet, little knowledge is available about the scope of alternative splicing at the whole genome level in normal and diseased hearts and even less about the mechanisms underlying the regulation of mRNA splicing in response to pathological injury in the heart. Recent studies have begun to fill this critical gap of information by establishing the total transcriptome, including RNA splicing variants, in normal and diseased hearts using RNA-seq and extensive bioinformatic, molecular, cellular, and functional analyses (Fig. 1).

4 Experimental approaches for total cardiac transcriptome analysis

The main tools necessary for total transcriptome studies include mRNA-seq using next generation high-throughput sequencing technology, exon assembly using bioinformatic tools, and the validation of exon boundaries and expression by quantitative reverse-transcription polymerase chain reaction (qRT-PCR), urea-polyacrylamide gel electrophoresis (UREA-PAGE), and capillary electrophoresis.

RNA-seq is a powerful high throughput sequencing technology involving the generation of a quantitative, genome scale, and single base resolution profile of the transcriptome (Anders *et al*., 2012; Li *et al*., 2012; Sanchez-Pla *et al*., 2012). The details of RNA-seq technology can be found in a recent review (Wang *et al*., 2009). In general, the mRNA is enriched from a sample of interest followed by the construction of a complementary DNA (cDNA) library using standard reverse-transcription methods. High throughput sequencing is performed using one of several technological platforms, including the illumina genome analyzer, Applied Biosystems (ABI) solid sequencing, and life science's 454 sequencing. This is an area of rapid improvement where speed, fidelity, and read lengths are increasing dramatically while the overall cost/base is dropping sharply. High throughput sequencing has become a routine method for scientific discovery and advanced clinical diagnosis. Its widespread application has already revolutionized our experimental

approaches where visualizing global changes and regulation in gene expression and transcriptome remodeling have become a reality. Indeed, RNA-seq has for the first time made it feasible to catalogue and appreciate the genome wide landscape of the whole transcriptome at single base resolution.

The output of the RNA-seq method is hundreds of millions of RNA sequence reads of about 70 to 100 bases in length. Linking these short reads into a contiguous transcript relies on a sophisticated computational algorithm. In general, the program first needs to map the reads on a particular exon based on matching sequences between the reads and genomic sequences, and then the exon-exon boundary is mapped based on a predicated cDNA database. Finally, all reads associated with a particular gene are combined to generate the total reads for each exon. Therefore, the final profile of each gene contains total reads of each exon at single base resolution. These mapping processes are complicated by a number of issues, including ultra-large datasets and limitations in computational power, repetitive sequences in closely related genes (miss-matching), incomplete genomic databases, sequencing errors, and sensitivity vs. fidelity (Mcintyre *et al.*, 2011; Ozsolak and Milos, 2011). Therefore, to generate a comprehensive expression profile for each exon, it is essential to perform the RNA-seq at sufficient depth. One major advantage of RNA-seq over micro-arrays is the possibility to identify novel, un-annotated exons or transcripts (Daines *et al.*, 2011; Lee *et al.*, 2011; Concha *et al.*, 2012). In a recent study, we have developed two bioinformatic tools, guided transcriptome reconstruction and de novo reconstruction. These tools allow detection of novel exons in known genes and novel transcript clusters (NTCs) (Lee *et al.*, 2011).

The quantification and specificity of identified known or novel exons should be validated by independent methods, including qRT-PCR. Different transcripts resulting from alternative RNA splicing can be separated by UREA-PAGE and capillary electrophoresis based on size differences. Given the fact that the reads generated from RNA-seq are assembled based on computational analysis, an experimental validation of the findings is always necessary. Indeed, we have identified and confirmed a significant number of differentially expressed exons in normal and diseased hearts by both fluorescent RT-PCR followed by UREA-PAGE and capillary electrophoresis, which show a very high correlation with the bioinformatics prediction (Fig. 2).

5 Global profiling of alternative RNA splicing in the heart: novel exons

In this study, deep RNA-seq was performed on mRNA samples prepared from adult mouse hearts in basal condition and in failing state as a result of chronic pressure-overload induced by trans-aortic constriction. This is a well-established model system to investigate the pathogenesis of cardiac hypertrophy and heart failure due to mechanical overload, mimicking chronic hypertension in humans.

Among the mRNA transcripts annotated from the RNA-seq data, more than 1000 novel exons were identified that had not been reported in any published databases. From a selected list of 40 novel exons, 38 (95%) were validated by RT-PCR in mouse heart tissue and all of them were further confirmed by direct DNA sequencing to have the predicted novel exon-exon junctions. The genes containing these novel exons included established regulators in cardiac signaling, mitochondria dynamics, and gene regulation. Many of the novel exons are predicted to have major functional impacts on the parent genes, including mRNA stability, protein truncation, protein activity, and post-translation modification. Using semi-qRT-PCR in human heart failure samples, some of these novel exons showed differential expression patterns in normal or diseased hearts, strongly suggesting that these novel exons may have a functional role in the disease. Therefore, deep RNA-seq revealed a significant number of novel transcripts, which contribute to overall transcriptome complexity in the heart.

6 Differential alternative splicing as a part of global transcriptome remodeling in developing and failing hearts

The onset of heart failure is associated with a significant change in both the quantity and quality of the cardiac transcriptome (Barry *et al.*, 2008;

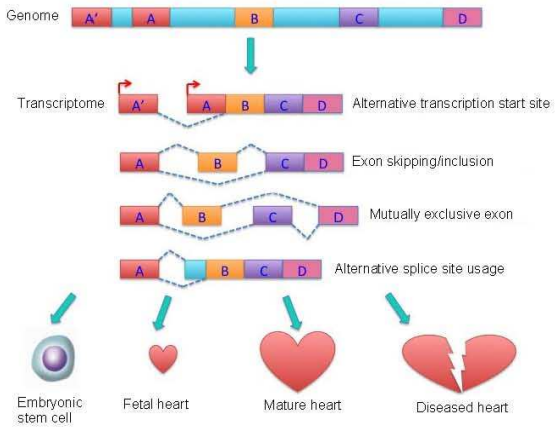


Fig. 1 Alternative RNA splicing contributing to the overall complexity of the cardiac transcriptome during cardiomyocyte differentiation, cardiac development, and response to pathological stress

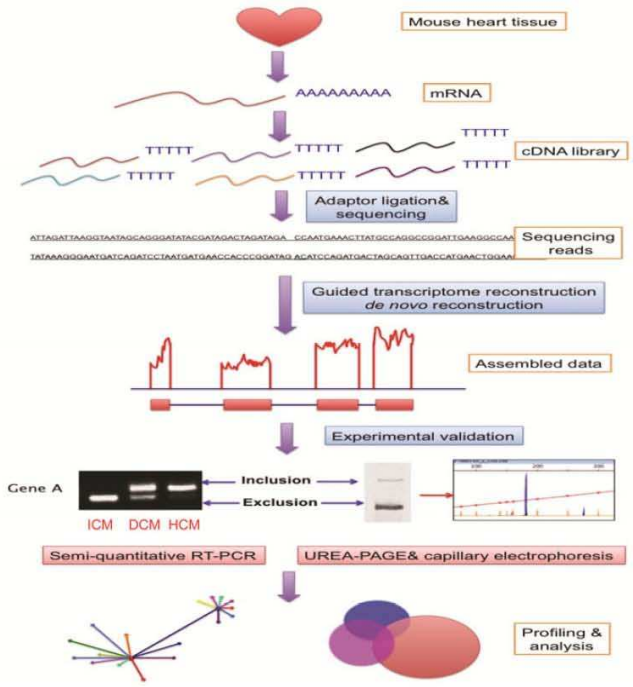


Fig. 2 Workflow of deep RNA-seq on profiling the cardiac transcriptome
ICM: ischemic cardiomyopathy; DCM: dilated cardiomyopathy; HCM: hypertrophic cardiomyopathy

Margulies *et al*., 2009). Most studies profiling the cardiac transcriptome have employed microarray approaches, which revealed global changes in gene expression in diseased hearts (Asakura and Kitakaze, 2009; Dewey *et al*., 2011). The features of transcriptome remodeling and the underlying regulatory mechanisms have been the focus of extensive investigation, leading to the identification of a network of responsible transcription factors and co-factors, including myocyte enhancer factor-2 (MEF-2), GATA, and histone deacetylases (HDACs) (Edmondson *et al*., 1994; Skerjanc *et al*., 1998; Naya *et al*., 2002; Backs and Olson, 2006). However, the global transcriptome profile of the alternatively spliced exons in cardiac development and disease remains to be established. RNA-seq analyses of cardiac transcriptome throughout heart development and disease progression hold great promise to address this question. Considering the potential impact of differentially expressed exons on protein function and regulation, alternative RNA splicing may emerge to be an important element in the underlying molecular mechanisms of cardiac lineage commitment, maturation, physiological or pathological responses to stresses. Studies on alternative RNA splicing events, the regulators and the functional consequences at the genome level will open a new frontier for us to explore the fundamental mechanisms of heart disease and potential therapeutic intervention.

References

Anders, S., Reyes, A., Huber, W., 2012. Detecting differential usage of exons from RNA-seq data. *Genome Res.*, in press [doi:10.1101/gr.133744.111]

Asakura, M., Kitakaze, M., 2009. Global gene expression profiling in the failing myocardium. *Circ. J.*, 73(9): 1568-1576.

Backs, J., Olson, E.N., 2006. Control of cardiac growth by histone acetylation/deacetylation. *Circ. Res.*, **98**(1):15-24. [doi:10.1161/01.RES.0000197782.21444.8f]

Barry, S.P., Davidson, S.M., Townsend, P.A., 2008. Molecular regulation of cardiac hypertrophy. *Int. J. Biochem. Cell Biol.*, **40**(10):2023-2039. [doi:10.1016/j.biocel.2008.02.020]

Bland, C.S., Wang, E.T., Vu, A., David, M.P., Castle, J.C., Johnson, J.M., Burge, C.B., Cooper, T.A., 2010. Global regulation of alternative splicing during myogenic differentiation. *Nucleic Acids Res.*, **38**(21):7651-7664. [doi:10.1093/nar/gkq614]

Buljan, M., Chalancon, G., Eustermann, S., Wagner, G.P., Fuxreiter, M., Bateman, A., Babu, M.M., 2012. Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell*, **46**(6):871-883. [doi:10.1016/j.molcel.2012.05.039]

Chen, M., Manley, J.L., 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.*, **10**(11):741-754. [doi:10.1038/nrm2777]

Concha, M., Wang, X., Cao, S., Baddoo, M., Fewell, C., Lin, Z., Hulme, W., Hedges, D., Mcbride, J., Flemington, E.K., 2012. Identification of new viral genes and transcript isoforms during epstein-barr virus reactivation using RNA-seq. *J. Virol.*, **86**(3):1458-1467. [doi:10.1128/jvi.06537-11]

Daines, B., Wang, H., Wang, L., Li, Y., Han, Y., Emmert, D., Gelbart, W., Wang, X., Li, W., Gibbs, R., *et al*., 2011. The drosophila melanogaster transcriptome by paired-end RNA sequencing. *Genome Res.*, **21**(2):315-324. [doi:10.1101/gr.107854.110]

Dasgupta, T., Ladd, A.N., 2012. The importance of celf control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *Wiley Interdiscip. Rev. RNA*, **3**(1):104-121. [doi:10.1002/wrna.107]

David, C.J., Chen, M., Assanah, M., Canoll, P., Manley, J.L., 2010. hnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature*, **463**(7279):364-368. [doi:10.1038/nature08697]

de la Grange, P., Gratadou, L., Delord, M., Dutertre, M., Auboeuf, D., 2010. Splicing factor and exon profiling across human tissues. *Nucleic Acids Res.*, **38**(9): 2825-2838. [doi:10.1093/nar/gkq008]

Dewey, F.E., Perez, M.V., Wheeler, M.T., Watt, C., Spin, J., Langfelder, P., Horvath, S., Hannenhalli, S., Cappola, T.P., Ashley, E.A., 2011. Gene coexpression network topology of cardiac development, hypertrophy and failure. *Circ. Cardiovasc. Genet.*, **4**(1):26-35. [doi:10.1161/CIRCGENETICS.110.941757]

Ding, J.H., Xu, X., Yang, D., Chu, P.H., Dalton, N.D., Ye, Z., Yeakley, J.M., Cheng, H., Xiao, R.P., Ross, J., *et al*., 2004. Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. *EMBO J.*, **23**(4):885-896. [doi:10.1038/sj.emboj.7600054]

Edmondson, D.G., Lyons, G.E., Martin, J.F., Olson, E.N., 1994. *Mef2* gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. *Development*, **120**(5):1251-1263.

Gang, H., Hai, Y., Dhingra, R., Gordon, J.W., Yurkova, N., Aviv, Y., Li, H., Aguilar, F., Marshall, A., Leygue, E., *et al*., 2011. A novel hypoxia-inducible spliced variant of mitochondrial death gene *Bnip3* promotes survival of ventricular myocytes. *Circ. Res.*, **108**(9):1084-1092. [doi:10.1161/circresaha.110.238709]

Guo, W., Schafer, S., Greaser, M.L., Radke, M.H., Liss, M., Govindarajan, T., Maatz, H., Schulz, H., Li, S., Parrish, A.M., *et al*., 2012. *RBM20*, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med.*, **18**(5):766-773. [doi:10.1038/nm.2693]

Hallegger, M., Llorian, M., Smith, C.W.J., 2010. Alternative splicing: global insights. *FEBS J.*, **277**(4):856-866.

[doi:10.1111/j.1742-4658.2009.07521.x]

Honda, A., Valogne, Y., Bou Nader, M., Brechot, C., Faivre, J., 2012. An intron-retaining splice variant of human cyclin A2, expressed in adult differentiated tissues, induces a g1/s cell cycle arrest in vitro. *PLoS One*, **7**(6):e39249. [doi:10.1371/journal.pone.0039249]

Kalsotra, A., Wang, K., Li, P.F., Cooper, T.A., 2010. MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development. *Genes Dev.*, **24**(7):653-658. [doi:10.1101/gad.1894310]

Koshelev, M., Sarma, S., Price, R.E., Wehrens, X.H.T., Cooper, T.A., 2010. Heart-specific overexpression of CUGBP1 reproduces functional and molecular abnormalities of myotonic dystrophy type 1. *Hum. Mol. Genet.*, **19**(6): 1066-1075. [doi:10.1093/hmg/ddp570]

Lee, J.H., Gao, C., Peng, G., Greer, C., Ren, S., Wang, Y., Xiao, X., 2011. Analysis of transcriptome complexity through RNA sequencing in normal and failing murine hearts. *Circ. Res.*, 109(12):1332-1341. [doi:10.1161/circresaha.111.249433]

Li, G., Bahn, J.H., Lee, J.H., Peng, G., Chen, Z., Nelson, S.F., Xiao, X., 2012. Identification of allele-specific alternative mRNA processing via transcriptome sequencing. *Nucleic Acids Res.*, in press [Epub ahead of print]. [doi:10.1093/nar/gks280]

Linke, W.A., Bucker, S., 2012. King of hearts: a splicing factor rules cardiac proteins. *Nat. Med.*, **18**(5):660-661. [doi:10.1038/nm.2762]

Margulies, K.B., Bednarik, D.P., Dries, D.L., 2009. Genomics, transcriptional profiling, and heart failure. *J. Am. Coll. Cardiol.*, **53**(19):1752-1759. [doi:10.1016/j.jacc.2008.12.064]

Mcintyre, L., Lopiano, K., Morse, A., Amin, V., Oberg, A., Young, L., Nuzhdin, S., 2011. RNA-seq: technical variability and sampling. *BMC Genomics*, **12**(1):293. [doi:10.1186/1471-2164-12-293]

Naya, F.J., Black, B.L., Wu, H., Bassel-Duby, R., Richardson, J.A., Hill, J.A., Olson, E.N., 2002. Mitochondrial deficiency and cardiac sudden death in mice lacking the MEF2A transcription factor. *Nat. Med.*, **8**(11):1303-1309. [doi:10.1038/nm789]

Ozsolak, F., Milos, P.M., 2011. RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.*, **12**(2): 87-98. [doi:10.1038/nrg2934]

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**(12):1413-1415. [doi:10.1038/ng.259]

Raghavachari, N., Barb, J., Yang, Y., Liu, P., Woodhouse, K., Levy, D., O'Donnell, C., Munson, P.J., Kato, G., 2012. A systematic comparison and evaluation of high density exon arrays and RNA-seq technology used to unravel the peripheral blood transcriptome of sickle cell disease. *BMC Med. Genomics*, **5**(1):28. [doi:10.1186/1755-8794-5-28]

Ramchatesingh, J., Zahler, A., Neugebauer, K., Roth, M., Cooper, T., 1995. A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.*, **15**(9):4898-4907.

Refaat, M.M., Lubitz, S.A., Makino, S., Islam, Z., Frangiskakis, J.M., Mehdi, H., Gutmann, R., Zhang, M.L., Bloom, H.L., Macrae, C.A., *et al.*, 2012. Genetic variation in the alternative splicing regulator *RBM20* is associated with dilated cardiomyopathy. *Heart Rhythm*, **9**(3):390-396. [doi:10.1016/j.hrthm.2011.10.016]

Sanchez-Pla, A., Reverter, F., Ruiz de Villa, M.C., Comabella, M., 2012. Transcriptomics: mRNA and alternative splicing. *J. Neuroimmunol.*, **248**(1-2):23-31. [doi:10.1016/j.jneuroim.2012.04.008]

Skerjanc, I.S., Petropoulos, H., Ridgeway, A.G., Wilton, S., 1998. Myocyte enhancer factor 2C and Nkx2-5 up-regulate each other's expression and initiate cardiomyogenesis in P19 cells. *J. Biol. Chem.*, **273**(52): 34904-34910. [doi:10.1074/jbc.273.52.34904]

Sultan, M., Schulz, M.H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., *et al.*, 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*, **321**(5891):956-960. [doi:10.1126/science.1160342]

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., Burge, C.B., 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**(7221):470-476. [doi:10.1038/nature07509]

Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.*, **10**(1):57-63. [doi:10.1038/nrg2484]

Warf, M.B., Berglund, J.A., 2007. MBNL binds similar RNA structures in the cug repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA*, **13**(12):2238-2251. [doi:10.1261/rna.610607]

Xu, X., Yang, D., Ding, J.H., Wang, W., Chu, P.H., Dalton, N.D., Wang, H.Y., Bermingham, J.R.Jr., Ye, Z., Liu, F., *et al.*, 2005. ASF/SF2-regulated CAMKIIδ alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell*, **120**(1):59-72. [doi:10.1016/j.cell.2004.11.036]

Yae, T., Tsuchihashi, K., Ishimoto, T., Motohara, T., Yoshikawa, M., Yoshida, G.J., Wada, T., Masuko, T., Mogushi, K., Tanaka, H., *et al.*, 2012. Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell. *Nat. Commun.*, **3**:883. [doi:10.1038/ncomms1892]

Chapter Five

RBFox1 Mediated Alternative RNA Splicing Regulates Development and Function in Heart

RBFox1 Mediated Alternative RNA Splicing Regulates Development and Function in Heart

Chen Gao[1,2,3], Shuxun Ren[1,2,3], Jae-Hyung Lee[1,5], Haodong Chen[1,2,3], Jared Churko[4], Joseph Wu[4],Thomas Vondriska[1,2,3], Preeti Ahuja[1,2,3], Xinshu Xiao[1,5], Jau-Nian Chen[1,2,6], Yibin Wang[1,2,3]

1. The Molecular Biology Institute, UCLA
2. Cardiovascular Research Laboratories, UCLA
3. Division of Molecular Medicine, Departments of Anesthesiology, Physiology and Medicine, David Geffen School of Medicine, UCLA
4. Departments of Medicine &Radiology, Stanford School of Medicine
5. Department of Integrative Biology and Physiology, UCLA
6. Department of Molecular and Cellular Developmental Biology, UCLA

Summary

RNA splicing contributes significantly to total transcriptome complexity but its functional role and regulation in cardiac development and diseases remain poorly understood. Based on total transcriptome analysis, we identified a significant number of alternative RNA splicing events in mouse failing hearts that resembled the pattern in fetal hearts. A muscle specific isoform of an RNA splicing regulator RBFox1 (A2BP1) is induced during cardiac development. Inactivation of zRBFox1 gene in zebrafish led to lethal phenotype associated with impaired cardiac function. RBFox1 regulates alternative splicing of transcription factor MEF2s, producing splicing variants with distinct transcriptional activities and different impact on cardiac development. RBFox1 expression is diminished in mouse and human failing hearts. Restoring RBFox1 expression significantly attenuates hypertrophy and heart failure induced by pressure-overload in mice. Therefore, RBFox1-MEF2 represents a previously uncharacterized regulatory circuit in cardiac transcriptional network with important impact on both cardiac development and diseases.

Introduction

Alternative RNA splicing significantly contributes to the total complexity of the transcriptome and helps to define the cellular identity with an estimated ~100, 000 intermediate to high abundant alternative splicing events in major human tissues (Pan, Shai et al. 2008). RNA splicing is a ubiquitous post-transcriptional process for all multi-exon genes in eukaryotes (Chen and Manley 2009). Alternative RNA splicing from individual genes can produce multiple mature mRNA species, yielding different protein isoforms with related, distinct or even opposing functions (Wang, Sandberg et al. 2008, Buljan, Chalancon et al. 2012). Alternative RNA splicing is regulated by cis-regulatory enhancers and silencers located within pre-mRNAs interacting with trans-acting splicing factors(de la Grange, Gratadou et al. 2010), including heterogeneous nuclear ribonucleoprotein (hnRNP) and serine-arginine rich proteins (SR proteins)(Pan, Shai et al. 2008, Sultan, Schulz et al. 2008, Chen and Manley 2009, Bland, Wang et al. 2010). Mis-regulated alternative RNA splicing events have a significant role in human diseases, affecting cellular processes from cell cycle to cell death (Hallegger, Llorian et al. 2010, Gang, Hai et al. 2011, Honda, Valogne et al. 2012, Nalini Raghavachari 2012, Yae, Tsuchihashi et al. 2012).

Alternative mRNA splicing has been associated with cardiac development and diseases, affecting structural genes such as cardiac troponin T or signaling molecules, such as CaM Kinase II (Ramchatesingh, Zahler et al. 1995, Ding, Xu

et al. 2004, Xu, Yang et al. 2005). Inactivation of splicing regulators SC35 or RBM20 led to dilated cardiomyopathy in mouse and rat (Guo, Schafer et al. 2012, Linke and Bucker 2012, Refaat, Lubitz et al. 2012). In addition, CUG-BP1 and ETR-like factors (CELF)/Bruno-like family of RNA binding proteins and muscleblind-like (MBNL) proteins have also been found to impact on both cardiac development and function(Warf and Berglund 2007, Kalsotra, Wang et al. 2010, Koshelev, Sarma et al. 2010, Dasgupta and Ladd 2012). However, relative to the extensive knowledge of transcriptional regulation, little is known about the global RNA splicing pattern and the functional impact of RNA splicing in heart during cardiac development and disease progression.

Induced fetal gene expression is a common feature observed in pathologically stressed heart and is serving as effective molecular biomarkers for heart diseases(Olson 2006, Barry, Davidson et al. 2008, Kuwahara, Nishikimi et al. 2012). This so called “fetal gene expression program” manifested in the diseased heart is dictated by a network of transcription factors, including many key players also implicated in cardiac differentiation and maturation, such as MEF2s and GATA4 (Lee, Gao et al. 2011). Although alternative RNA splicing events have been widely observed among many cardiac genes, changes of alternative RNA splicing observed in the failing hearts have not been noted to have any significant association with the changes observed during cardiac development (Gao and Dudley Jr , Guo, Schafer et al. 2012). Thus, there is no evidence indicating that alternative RNA splicing regulation in cardiac diseases and development are related at mechanistically or functional levels.

Using deep RNA-seq approach, we characterized the cardiac transcriptome at single exon resolution from RNA samples prepared from adult mouse hearts at basal and in trans-aortic constriction (TAC)-induced failing state(Lee, Gao et al. 2011). By comparing the splicing pattern in failing vs. neonatal hearts, we found a fetal-like RNA splicing pattern in diseased hearts, affecting many cardiac genes including all members of the MEF2 family transcription factors. We discovered a muscle specific isoform of the splicing regulator RBFox1 (A2BP1) as a direct trans-acting regulator for MEF2 alternative splicing. RBFox1 was significantly induced during cardiac development and inactivation of RBFox1 in zebrafish led to abnormal MEF2 alternative splicing and impaired cardiac development and function. We further demonstrated that MEF2 isoforms resulted from RBFox1 mediated alternative RNA splicing possessed distinct transcriptional activities and *in vivo* function during development. Finally, loss of RBFox1 was observed in mouse and human failing hearts, and restoring RBFox1 expression in vivo preserved cardiac function and suppressed cardiac hypertrophy in response to pressure-overload. Therefore, our study has established RBFox1/MEF2 as an uncharacterized common molecular mechanism of RNA splicing regulation in both cardiac development and diseases. This novel regulatory circuit has a significant impact on transcriptome maturation and pathological re-programming in heart.

Results:

A Common Pattern of Alternative RNA Splicing Associated with Cardiac Postnatal Development and Heart Failure

Alternative RNA splicing is a common phenomenon observed in many cardiac genes during development or disease progression (Gao and Dudley Jr , Guo, Schafer et al. 2012). Using deep RNA-seq, we have identified transcriptome-wide alternative splicing events associated with pressure-overload induced heart failure in mice(Lee, Gao et al. 2011). Among the exons being alternatively spliced, we compared their relative expression by real-time RT-PCR in the ventricular samples of one day old neonates, 5 month old adult sham operated mice and age matched mice 8 weeks after trans-aortic constriction(Lee, Gao et al. 2011) (**Supplemental Figure 5.1**). From all 30 selected exons (**Supplemental Table 1, Supplemental Figure 5.2**), changes of exon utilization observed in the adult failing hearts resembled the pattern observed in the neonatal hearts (**Figure 5.1A**). This suggests that the alternative RNA splicing pattern observed in the adult failing heart shares characteristic features of a fetal heart, and this "fetal-like" reprogramming in RNA splicing is also part of the transcriptome changes following pathological stress.

The genes affected by alternative RNA splicing include all members of the MEF2 family, Mef2a, Mef2c and Mef2d, transcription factors implicated in both cardiac development and pathological reprogramming (Lin, Schwarz et al. 1997, Lu, McKinsey et al. 2000). As shown in **Figure 5.1B**, MEF2s share a common gene

structure where one of the two exons (a vs. b) 3' to the MADS/MEF2 domain coding sequences is selectively included in the mature transcripts due to a mutually exclusive RNA splicing event(Bachinski, Sirito et al. 2010). The relative expression ratio between exon b vs. exon a containing MEF2s was significantly induced during cardiac development in mouse and human (**Figure 5.1C-D**), and this ratio was also significant reversed in both mouse and human failing hearts (**Figure 5.1C-D**). Therefore, the fetal-like RNA splicing reprogramming is a conserved process in diseased hearts affecting many genes, including all members of the MEF2 family.

RBFox1 is a Candidate Trans-acting RNA Splicing Regulator in Cardiac Development and Heart Failure

The coordinated changes in RNA splicing patterns are mediated by cis-and trans-acting splicing regulatory elements (McManus and Graveley 2011, Witten and Ule 2011). In search for the putative trans-acting factors participating the fetal-like RNA splicing regulation, we performed a *de novo* motif discovery analysis on all the differentially spliced exons observed in failing mouse hearts. By screening for significantly enriched and evolutionarily conserved 5-mer binding motifs in the flanking introns and exonic regions of the affected exons (Xiao, Wang et al. 2007) we identified several binding motifs significantly enriched near differentially spliced exons (**Table 5.1**). One conserved binding sequence significantly enriched was GCATG/TGCAT for RBFox1 (a splicing regulator with specific expression pattern in brain and striated muscle (Jin, Suzuki et al.

2003). In parallel, we measured the expression levels of a panel of 25 known RNA splicing regulators in normal and pressure-overload induced failing hearts (**Supplement Figure 5.3**). Among them, only RBFox1 was significantly down-regulated in the failing hearts comparing to the Sham operated controls at both mRNA and protein level (**Figure 5.2A-B**). These data suggest that RBFox1 is a candidate trans-acting regulator for alternative RNA splicing in failing heart. RBFox1 expression was induced in heart during cardiac maturation in zebrafish, mouse and human (**Figure 5.2C-F, Supplemental Figure 5.5**). Therefore, RBFox1 has a highly conserved expression pattern during cardiac development and diseases, correlating well with the observed changes in RNA alternative splicing.

RBFox1 is Essential for Cardiac Development and Function in Zebrafish

To investigate the function of RBFox1 in heart, we examined the impact of RBFox1 inactivation in developing zebrafish embryos. As shown in Figure 3, RBFox1 morphants developed severe cardiac phenotype characterized by reduced ejection fraction, lower heart rate (**Figure 5.3J, K**) and presence of pericardial edema (**Figure 5.3 A,B**), along with collapsed ventricle, enlarged atria (**Figure 5.3D,E, Supplemental Movie 1,2**) as well as defects in circulation (**Figure 5.3L, Supplemental movie 4,5**). Injecting zebrafish RBFox1 mRNA into the morphants significantly rescued ventricular defects (**Figure 5.3I, Movie 3**) and attenuated the extent of cardiac dysfunction in the RBFox1 morphants at 48hpf (**Figure 5.3J,K**). Remarkably, overexpression of mouse RBFox1 mRNA also

significantly rescued the cardiac defects observed in the RBFox1 morphants (**Supplemental Figure 5.6**), suggesting a conserved role of RBFox1 in heart development. Finally, in situ hybridization on embryos at 54hpf using cardiac chamber specific markers, including amhc, vmhc and Notch1b, showed less-defined expression pattern of chamber specific cardiac genes in the RBFox1 morphants(**Figure 5.3M**). All these data suggests that RBFox1 is a functionally conserved RNA splicing regulator essential to normal cardiac development and function.

RBFox1 is a Necessary and Sufficient Trans-Acting Regulator for MEF2 Splicing in Heart

MEF2s are transcription factor family implicated in both cardiac development and pathological remodeling (Lin, Schwarz et al. 1997, Lu, McKinsey et al. 2000). Since MEF2 family members show a conserved alternative splicing pattern as part of the "fetal like" RNA splicing reprogramming in heart (**Figure 5.1C**), we investigated whether MEF2s are potential downstream targets of RBFox1. Indeed, a consensus RBFox1 binding motif was identified near the alternatively spliced exons for all MEF2 genes in fish, mouse and human genome (**Figure 5.4A**). In addition, changes in MEF2 isoform expression ratio was highly correlated with RBFox1 expression during cardiac development and heart failure (**Figure 5.4B**). To demonstrate the direct impact of RBFox1 expression on MEF2 alternative RNA splicing, mouse RBFox1 was expressed in neonatal rat ventricular myocytes (NRVM) via an adenoviral vector where the endogenous RBFox1 expression was

low (**Figure 5.4C**). RBFox1 expression in neonatal myocytes led to significant induction in the inclusion of MEF2 Exon b to Exon a ratio (**Figure 5.4D,E**). In contrast, using morpholino to inactivate RBFox1 in zebrafish embryo, we observed the cardiac MEF2 b vs. a exon expression ratios were significantly reduced (**Figure 5.4F**). To further establish MEF2 genes as direct downstream targets of RBFox1 mediated RNA splicing, we utilized an *in vitro* minigene reporter system(Boutz, Stoilov et al. 2007), including mouse Mef2a Exon 5b with the flanking intronic fragment containing the putative RBFox1 binding motif (**Figure 5.4G**). Upon co-expressing the reporter with RBFox1 in HEK293 cells, we detected increased Mef2a Exon 5b inclusion (**Figure 5.4G**). This induction was abolished when the putative RBFox1 binding site in the Mef2a intron was either mutated or deleted. Similar results were observed using a similar minigene reporter for the mouse Mef2d Exon 4b (**Supplemental Figure 5.6**). Combining these with the gain and loss of function studies described earlier, we conclude that RBFox1 is both a necessary and a sufficient trans-acting factor for the conserved alternative splicing of MEF2 genes in heart.

Regulatory Circuit of RBFox1-MEF2 in Heart

The remarkable conservation of the RBFox1-MEF2s circuit was manifested in correlated expression and the conserved presence of the RBFox1 binding motifs in the MEF2 genes among vertebrates. However, no prior study has demonstrated the functional differences between the two MEF2 isoforms expressing a vs. b exons. We first tested the functional impact of Mef2a-Exon 4a vs Mef2a-Exon 4b

isoforms in zebrafish embryos by individually over-expressing these two isoforms in developing zebrafish embryo. As shown in Figure 5A-C, embryos receiving synthetic Mef2a-E4a RNA displayed severe heart failure phenotype comparing to the control embryos, while zebrafish embryos overexpressing the Mef2a-E4b gene showed normal cardiac phenotype at comparable expression levels. Similarly, lethal cardiac phenotype was observed in embryos receiving synthetic mouse Mef2a-E5a isoform RNA at comparable level (**Figure 5.5D-E**). This data suggests that different Mef2a splicing variants have distinct function.

In addition, we simultaneously knocked down RBFox1 and individual Mef2a splicing variants using targeted morpholino (**Figure 5.6A, Supplemental Figure 5.7**). As shown in Figure 3, RBFox1 inactivation in zebrafish led to reduced inclusion of Exon 4b in Mef2a, leading to a significant induction of Mef2a exon 4a vs.4b expression ratio. Double knockdown of RBFox1 and Mef2a-E4a isoform significantly rescued the cardiac defects in the RBFox1 morphants as measured from the presence of pericardial edema, embryos viability and normal circulation (**Figure 5.6B-C**). In contrast, knockdown of RBFox1 in combination with knockdown of Mef2a-E4b failed to rescue the lethal cardiac phenotype comparing to RBFox1 morphant (**Figure 5.6A-B**). Therefore, both gain and loss of function studies of Mef2a isoforms clearly demonstrate that these splicing variants have distinct function in cardiac development. These in vivo evidences suggest that RBFox1 mediated regulation of the expression ratio of the MEF2 splicing variants is critical to normal cardiac development and function.

Mef2a Splicing Variants Direct Distinct Transcriptional Activity

Our in vivo observation in zebrafish contradicts with an earlier study showing different Mef2c isoforms generated by this alternative splicing event have similar transcriptional activity based on in vitro luciferase reporter assay(Bachinski, Sirito et al. 2010). However, the transcriptional activities of MEF2 splicing variants have not been studied for the endogenous downstream targets. To investigate that, we performed RNA-seq studies in Mef2a-Exon-4a and Mef2a-Exon-4b overexpressing zebrafish embryos at 24hpf (**Figure 5.7A**) at which stage the deleterious cardiac phenotype has not become apparent. Comparing to the control embryos, we identified a total of 2334 number of genes that were differentially expressed in the embryos expressing the two Mef2a isoforms. Among them, 905 genes were preferentially induced by Mef2a-E4b isoform and 209 genes were preferentially induced by Mef2a-E4a isoform. Most notably, several well established cardiac transcription factors, including Nkx2.5, Gata4, Tbx20 and Mef2d were preferentially induced by Mef2a-Exon-4b vs. 4a isoform (**Figure 5.7B**). To validate these observations in zebrafish, we specifically expressed different mouse Mef2a isoforms in NRVM via adenovirus vectors. Consistent with the results from zebrafish, Mef2a-E5b expression led to significantly higher induction of cardiac transcription factors, including Nkx2.5, Gata4, Tbx20 and Mef2d comparing to Mef2a-E5a splicing variant. Furthermore, using Nkx2.5 promoter driving luciferase reporter, we showed Mef2a-E5b has significantly higher activity to induce Nkx2.5 transcription than Mef2a-Exon-5a (**Figure 5.7E**). Similar observation was made for Tbx20 promoter (**Supplemental Figure 5.8**).

Therefore, RBFox1 mediated alternative splicing of MEF2 yielded two isoforms with distinct transcriptional activities towards downstream genes including some key cardiac transcription factors. These data offers a molecular basis for the essential role of RBFox1 and the differential impact of MEF2 isoforms in cardiac development and function.

RBFox1 Mediated RNA splicing in Cardiac Hypertrophy and Heart Failure

RBFox1 expression was low in neonatal hearts, elevated in adult hearts and significantly reduced again in mouse failing hearts. Ectopic expression of RBFox1 in NRVM promoted cell size growth (**Figure 5.8A**) but a significant suppression of "fetal gene" expression (**Figure 5.8B**). More importantly, significant RBFox1 down-regulation was observed in human dilated cardiomyopathy hearts (**Figure 5.8C**) and TAC induced RBFox1 down-regulation was reversed one day after pressure-overload was removed in mice along with some heart failure marker genes (**Figure 5.8D and Supplement Figure 5.4**). In contrast, modest induction of RBFox1 in adult mouse hearts in a cardiac specific RBFox1 transgenic model did not cause detectable basal phenotype with normal cardiac function (**Figure 5.8E, F**). However, following 6 weeks of pressure-overload induced by TAC, RBFox1 transgenic mice showed significantly preserved function comparing to non-transgenic littermate controls (**Figure 5.8G,H**), and significantly attenuated hypertrophy and heart failure gene expression (**Figure 5.8I**). Therefore, restoring RBFox1 expression significantly

blunted the deleterious pathological remodeling in stressed hearts at functional and molecular levels.

Discussion:

In this report, we characterized genome-wide changes in RNA splicing in heart during postnatal development and disease progression. From our RNA-seq and extensive validation studies, we uncovered a "fetal-like" alternative RNA splicing program in failing heart for a significant number of transcripts with important function in heart, including a highly conserved alternative splicing event for all members of the MEF2 gene family. Based on both *in vitro* and *in vivo* evidence, we further established that this alternative splicing of MEF2 is regulated by a tissue-specific splicing regulator--RBFox1. Using zebrafish and cultured myocytes, we demonstrated that RBFox1 mediated alternative splicing resulted in two MEF2 isoforms with distinct downstream gene activation profiles. At functional level, inactivation of RBFox1 caused cardiac dysfunction associated with defects in myocyte maturation that was significantly contributed by specific induction of the Mef2a splicing defects. Finally, loss of RBFox1 is associated with the onset of heart failure following pressure-overload while restoring RBFox1 expression in a transgenic model ameliorated the development of heart failure. Therefore, our study revealed a novel regulatory circuit in transcriptome programming and reprogramming during cardiac development and disease

involving RBFox1 mediated alternative splicing of cardiac genes, including MEF2s.

A Global Alternative Splicing Network Associates with Both Cardiac Development and Pathological Remodeling

In this study, we identified that the disease associated alternative RNA splicing pattern resembled what was observed in fetal hearts. Although a "fetal-like" transcriptional reprogramming has long been established in diseased hearts, a "fetal-like" RNA splicing program has not been observed or reported. Although we cannot confirm all RNA splicing events in failing hearts are "fetal like", the significant number of genes we have analyzed so far support the concept that a "fetal-like" RNA splicing reprogramming is wide-spread in diseased heart. A global and coordinated regulation of RNA splicing in both cardiac development and diseases implies a potentially common regulatory mechanism. However, comparing to the abundant knowledge of transcriptional regulation, very limited insights are available to cardiac RNA splicing regulation during development or diseases. Inactivation of splicing factors, SC35 and RBM20, has been found to cause dilated cardiomyopathy in mouse and rat(Guo, Schafer et al. 2012, Linke and Bucker 2012, Refaat, Lubitz et al. 2012). Other splicing factors shown to play important roles in cardiac development and function include CUGBP1 and ETR-likefactors, CELF/Bruno-like family of RNA binding proteins, and muscleblind-like (MBNL) proteins(Warf and Berglund 2007, Kalsotra, Wang et al. 2010, Koshelev, Sarma et al. 2010, Dasgupta and Ladd 2012).However, these splicing

regulators do not show tissue specific and dynamic expression pattern related to cardiac development and diseases. Identification of RBFox1 revealed the first tissue-specific RNA splicing regulator with an expression pattern significantly correlated with the changes of RNA splicing in hearts during development and pathogenesis. Our extensive in vitro and in vivo evidence further established that RBFox1 is a necessary and sufficient trans-acting regulator for MEF2 alternative splicing. RBFox1 expression is critical to normal cardiac development and has a significant impact on the development of cardiac hypertrophy and heart failure under pathological conditions. Therefore, defect of RBFox1 mediated RNA splicing is a newly established molecular component in diseased hearts. However the regulatory mechanism for RNA splicing in general, and for RBFox1 expression and function during heart failure in particular, is poorly understood and should be an interesting area for further studies.

Alternative Splicing Mediated Transcriptional Regulation

It is known that alternative RNA splicing affects many genes in heart, including genes encoding structural proteins, such as cardiac troponin T, or signaling molecules, such as Ca^{2+}/calmodulin-dependent protein kinase (CaM kinase) (Ramchatesingh, Zahler et al. 1995, Ding, Xu et al. 2004, Xu, Yang et al. 2005). The mutually exclusive alternative splicing of the MEF2 gene for an exon 3' adjacent to the MEF2 and DNA binding MADS box domains (encode MADS/MEF2 domain) has been reported as a highly conserved splicing event observed in species across vertebrates and shared by all MEF2 family members.

Yet, the functional significance of this MEF2 alternative splicing event is not known. The domain encoded by the two mutually exclusive exons was never included in structural studies on MEF2 (Wu, Dey et al. 2010). The regulation of this splicing event has never been studied, neither. Our study established for the first time that RBFox1 is a potent trans-acting regulator of the alternative splicing for MEF2 genes, and the expression of the MEF2 splicing variants changes dynamically in developing and diseased hearts, following a "fetal-like" RNA splicing pattern. In addition, with specific gain and loss of function studies, we demonstrated that the individual MEF2 splicing variants had significant differences in transcriptional targets and functional impact on cardiac development, maturation and pathological remodeling. These studies established for the first time that a coordinated regulation of MEF2 alternative splicing is a result of RBFox1 mediated RNA splicing and the splicing variants of MEF2 have distinct transcriptional activities that have a significant functional impact on normal cardiac development and diseases.

A Fine-tuning Transcriptional & Post-Transcriptional Network Regulating Cardiomyocytes Differentiation and Maturation

Both genetics analyses and molecular studies have demonstrated a central role for transcriptional regulation in cardiomyocyte lineage determination and continuing maturation, involving key transcription factors and genome-wide chromatin remodeling (Qian, Huang et al. 2012, Song, Nam et al. 2012). Our study has revealed a new dimension in the transcriptome programming network involving

RBFox1 mediated RNA splicing of transcription factors, including MEF2. Although not an essential pathway for cardiomyocyte lineage determination, RBFox1 expression is necessary for normal cardiac development and function beyond the initial commitment. The fact that RBFox1 expression is further induced in adult hearts comparing to the fetal or neonatal heart suggests that RBFox1 mediated RNA splicing plays a necessary role to refine the cardiac transcriptome into its matured form. Since MEF2 splicing variants have distinct downstream transcriptional targets, in addition to their total level of expression, the relative ratio of the MEF2 splicing variants can determine and refine the transcriptome composition in heart. Therefore, it is plausible that RBFox1-MEF2 pathway is a fine-tuning regulatory circuit for cardiac transcriptome maturation in adult hearts. Finally, our in vivo studies showed that restoring RBFox1 expression in mice attenuated pressure-overload induced heart failure and re-balancing MEF2 splicing variants ratio in zebrafish embryo also rescued development defects caused by RBFox1 inactivation. Therefore, RBFox1-MEF2 regulatory circuit has a significant contribution to cardiac development and disease progression, and can serve as a new target of therapeutic intervention.

EXPERIMENTAL PROCEDURES:

Zebrafish

Adult zebrafish and embryos were maintained as previously described(Westerfield 1995). Embryos for in situ hybridization were raised in the presence of 0.2mM 1-phenyl-2-thiourea to maintain optical transparency(Westerfield 1995). The *Cmlc:GFP* strain used for this study has been previously described (Lu, Ren et al. 2007).

RBFox1 Transgenic Mouse:

Animals in this study were handled in accordance with the *Guide for the Care and Use of Laboratory Animals published by the US National Institutes of Health*. The detailed description of the generation of RBFox1 transgenic mice is provided in Supplemental Material.

Zebrafish Embryo Morpholino and mRNA injection

Morpholino antisense oligonucleotides (MOs) complementary to the translation start site and its flanking sequence for zebrafish RBFox1 gene and mRNAs for Mef2a isforms were used with details described in Supplemental Material.

Tissue from human Non-failing and Failing Hearts

The failing heart samples (n=16) were obtained from the left ventricular (LV) anterior wall during heart transplantation or implantation of an LV assist device

(Ref). The non-failing heart samples (NF) (n=8) were obtained from the LV free wall and procured from National Disease Research Interchange (NDRI) and University of Pennsylvania. NF heart donors had no history of macroscopic or laboratory signs of cardiac diseases. The tissue collection was approved by the UCLA Institutional Review Board #11-001053 and #12-000207. Fetal heart tissue was procured from medical waste (StemExpress, Diamond Springs, CA) as approved by the Stanford Institutional Review Board.

Pressure-overload model of HF in mouse

Left ventricle tissues were collected from male C57BL/6 mice 8 weeks post trans-aortic constriction (TAC) procedure (HF) and 1 day post birth (Neonatal) respectively and their corresponding Sham controls as described(Lee, Gao et al. 2011). Doppler velocity measurement of right and left carotid arteries were obtained from TAC treated mice to confirm the consistency of the surgery procedure. The heart failure status of the TAC treated animals was established based on a significant increase in heart weight and a significant reduction in ejection fractions measured by echocardiography.

Gene expression analysis

Gene expression analysis via RNA-seq, qRT-PCR, in-situ hybridization, luciferase reporter gene assays were described in details in Supplemental Material.

Statistical Analysis

Data are expressed as mean ± STDEV. For comparison between two groups, differences were analyzed by Student's t-test. For multiple groups' comparison, differences were analyzed by one-way ANOVA.p values ≤ 0.05 were considered as significant.

ACKNOWLEGEMENT:

Authors wish to thank Ms. Haiying Pu for her technical assistance. We thank Dr. Douglas Black for reagents and discussion. This work was supported in part by grants from National Institutes of Health (HL070079, HL103205, HL108186 and HL110667 for YW, HL096980 for JC, UCLA CTSI-Cardiovascular Pilot Team Research Grant UL1TR000124 (YW, XX, TV), American Heart Association Established Investigator Award (YW), Eli &Edythe Broad Center Pre-Doctoral Fellowship in Stem Cell Science (CG), and UCLA Cardiovascular Research Laboratories.

REFERENCES:

Adamidi, C. (2011). "De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics." Genome Res. **21**: 1193-1200.

Au, K. F., H. Jiang, L. Lin, Y. Xing and W. H. Wong (2010). "Detection of splice junctions from paired-end RNA-seq data by SpliceMap." Nucleic Acids Res. **38**: 4570-4578.

Bachinski, L. L., M. Sirito, M. Böhme, K. A. Baggerly, B. Udd and R. Krahe (2010). "Altered MEF2 isoforms in myotonic dystrophy and other neuromuscular disorders." Muscle & Nerve **42**(6): 856-863.

Barry, S. P., S. M. Davidson and P. A. Townsend (2008). "Molecular regulation of cardiac hypertrophy." Int J Biochem Cell Biol **40**(10): 2023-2039.

Barry, S. P., S. M. Davidson and P. A. Townsend (2008). "Molecular regulation of cardiac hypertrophy." The International Journal of Biochemistry & Cell Biology **40**(10): 2023-2039.

Birol, I. (2009). "De novo transcriptome assembly with ABySS." Bioinformatics **25**: 2872-2877.

Bland, C. S., E. T. Wang, A. Vu, M. P. David, J. C. Castle, J. M. Johnson, C. B. Burge and T. A. Cooper (2010). "Global regulation of alternative splicing during myogenic differentiation." Nucleic Acids Res **38**(21): 7651-7664.

Boutz, P. L., P. Stoilov, Q. Li, C.-H. Lin, G. Chawla, K. Ostrow, L. Shiue, M. Ares and D. L. Black (2007). "A post-transcriptional regulatory switch in

polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons." Genes & Development **21**(13): 1636-1652.

Buljan, M., G. Chalancon, S. Eustermann, G. P. Wagner, M. Fuxreiter, A. Bateman and M. M. Babu (2012). "Tissue-Specific Splicing of Disordered Segments that Embed Binding Motifs Rewires Protein Interaction Networks." Molecular Cell **46**(6): 871-883.

Chang, S., T. A. McKinsey, C. L. Zhang, J. A. Richardson, J. A. Hill and E. N. Olson (2004). "Histone Deacetylases 5 and 9 Govern Responsiveness of the Heart to a Subset of Stress Signals and Play Redundant Roles in Heart Development." Molecular and Cellular Biology **24**(19): 8467-8476.

Chen Gao, Y. W. (2012). "Global impact of RNA splicing on transcriptome remodeling in the heart." Journal of Zhejiang University Science B **13**(8): 603-608.

Chen, M. and J. L. Manley (2009). "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." Nat Rev Mol Cell Biol **10**(11): 741-754.

Dasgupta, T. and A. N. Ladd (2012). "The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins." Wiley Interdisciplinary Reviews: RNA **3**(1): 104-121.

de la Grange, P., L. Gratadou, M. Delord, M. Dutertre and D. Auboeuf (2010). "Splicing factor and exon profiling across human tissues." Nucleic Acids Research **38**(9): 2825-2838.

Deepak Srivastava, E. N. O. (2000). "A genetic blueprint for cardiac development." Nature **407**: 221-226.

Ding, J.-H., X. Xu, D. Yang, P.-H. Chu, N. D. Dalton, Z. Ye, J. M. Yeakley, H. Cheng, R.-P. Xiao, J. Ross, J. Chen and X.-D. Fu (2004). "Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart." EMBO J **23**(4): 885-896.

Durocher, D., F. Charron, R. Warren, R. J. Schwartz and M. Nemer (1997). "The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors." EMBO J. **16**: 5687-5696.

Edwards, A. S. and A. C. Newton (1997). "Phosphorylation at Conserved Carboxyl-terminal Hydrophobic Motif Regulates the Catalytic and Regulatory Domains of Protein Kinase C." Journal of Biological Chemistry **272**(29): 18382-18390.

Eric Chruchill, G. B., Alice Vallentin, Tomoyoshi Koyanagi, and Daria Mochly-Rosen (2008). "PKC Isozymes in Chronic Cardiac Disease: Possible Therapeutic Targets." Annual Reviews Pharmacology and Toxicology(48): 569-599.

Gajewski, K., Y. Kim, Y. M. Lee, E. N. Olson and R. A. Schulz (1998). "D-Mef2: a target for tinman activation during Drosophila heart development." EMBO J. **16**: 515-522.

Gang, H., Y. Hai, R. Dhingra, J. W. Gordon, N. Yurkova, Y. Aviv, H. Li, F. Aguilar, A. Marshall, E. Leygue and L. A. Kirshenbaum (2011). "A Novel Hypoxia-Inducible Spliced Variant of Mitochondrial Death Gene Bnip3 Promotes

Survival of Ventricular Myocytes / Novelty and Significance." Circulation Research **108**(9): 1084-1092.

Gao, G. and S. C. Dudley Jr "RBM25/LUC7L3 Function in Cardiac Sodium Channel Splicing Regulation of Human Heart Failure." Trends in Cardiovascular Medicine(0).

Gao, G. and S. C. Dudley Jr (2013). "RBM25/LUC7L3 Function in Cardiac Sodium Channel Splicing Regulation of Human Heart Failure." Trends in Cardiovascular Medicine **23**(1): 5-8.

Greulich, F., C. Rudat and A. Kispert (2011). "Mechanisms of T-box gene function in the developing heart." Cardiovascular Research **91**(2): 212-222.

Guo, W., S. Schafer, M. L. Greaser, M. H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A. M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. A. Hacker, K. W. Saupe, G. W. Dec, P. T. Ellinor, C. A. MacRae, B. Spallek, R. Fischer, A. Perrot, C. Ozcelik, K. Saar, N. Hubner and M. Gotthardt (2012). "RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing." Nat Med **18**(5): 766-773.

Gusterson, R. J., E. Jazrawi, I. M. Adcock and D. S. Latchman (2003). "The Transcriptional Co-activators CREB-binding Protein (CBP) and p300 Play a Critical Role in Cardiac Hypertrophy That Is Dependent on Their Histone Acetyltransferase Activity." Journal of Biological Chemistry **278**(9): 6838-6847.

Hallegger, M., M. Llorian and C. W. J. Smith (2010). "Alternative splicing: global insights." FEBS Journal **277**(4): 856-866.

Hambleton, M., H. Hahn, S. T. Pleger, M. C. Kuhn, R. Klevitsky, A. N. Carr, T. F. Kimball, T. E. Hewett, G. W. Dorn, W. J. Koch and J. D. Molkentin (2006). "Pharmacological- and Gene Therapy-Based Inhibition of Protein Kinase Cα/β Enhances Cardiac Contractility and Attenuates Heart Failure." Circulation **114**(6): 574-582.

Harvey, R. P. (1996). Dev. Biol. **178**: 203-216.

Henderson, D. J. and A. J. Copp (1998). "Versican Expression Is Associated With Chamber Specification, Septation, and Valvulogenesis in the Developing Mouse Heart." Circulation Research **83**(5): 523-532.

Honda, A., Y. Valogne, M. Bou Nader, C. Br√©chot and J. Faivre (2012). "An Intron-Retaining Splice Variant of Human Cyclin A2, Expressed in Adult Differentiated Tissues, Induces a G1/S Cell Cycle Arrest In Vitro." PLoS ONE **7**(6): e39249.

Hong Ki Song, S.-E. H., Taeyong Kim, Do Han Kim (2012). "Deep RNA Sequencing Reveals Novel Cardiac Transcriptomic Signatures for Physiological and Pathological Hypertrophy." PLOS ONE.

Huang, Z.-P., J. Chen, H. Y. Seok, Z. Zhang, M. Kataoka, X. Hu and D.-Z. Wang (2013). "MicroRNA-22 Regulates Cardiac Hypertrophy and Remodeling in Response to Stress." Circulation Research **112**(9): 1234-1243.

Jin, Y., H. Suzuki, S. Maegawa, H. Endo, S. Sugano, K. Hashimoto, K. Yasuda and K. Inoue (2003). "A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG." EMBO J **22**(4): 905-912.

Jingjing, L., Y. Xue, N. Agarwal and R. S. Roque (1999). "Human Müller cells express VEGF183, a novel spliced variant of vascular endothelial growth factor." Investigative Ophthalmology & Visual Science **40**(3): 752-759.

Joanna Kobylinska, W. D., Monika Cendrowska-Pinkosz, Anna Dworzanska, Teresa Hermanowics-Dryka, Joanna Kiszka, Elzbieta Staroslawska, Franciszek Burdan (2013). "Morphological and Moelcular Bases of cardiac development." Postepy Hig Med Dosw.

Jr, G. G. a. S. C. D. (2013). "SCN5A splicing variants and the possibility of predicting heart failure-associated arrhythmia." Expert Reveiw of Cardiovascular Therapy **11**(2): 117-119.

Kalsotra, A., K. Wang, P.-F. Li and T. A. Cooper (2010). "MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development." Genes & Development **24**(7): 653-658.

Kilic, A., A. Velic, L. J. De Windt, L. Fabritz, M. Voss, D. Mitko, M. Zwiener, H. A. Baba, M. van Eickels, E. Schlatter and M. Kuhn (2005). "Enhanced Activity of the Myocardial Na+/H+ Exchanger NHE-1 Contributes to Cardiac Remodeling in Atrial Natriuretic Peptide Receptor–Deficient Mice." Circulation **112**(15): 2307-2317.

Koshelev, M., S. Sarma, R. E. Price, X. H. T. Wehrens and T. A. Cooper (2010). "Heart-specific overexpression of CUGBP1 reproduces functional and molecular abnormalities of myotonic dystrophy type 1." Human Molecular Genetics **19**(6): 1066-1075.

Kuwahara, K., T. Nishikimi and K. Nakao (2012). "Transcriptional Regulation of the Fetal Cardiac Gene Program." Journal of Pharmacological Sciences **119**(3): 198-203.

Kuwahara, K., T. Nishikimi and K. Nakao (2012). "Transcriptional regulation of the fetal cardiac gene program." J Pharmacol Sci **119**(3): 198-203.

Ladage, D., L. Tilemann, K. Ishikawa, R. N. Correll, Y. Kawase, S. R. Houser, J. D. Molkentin and R. J. Hajjar (2011). "Inhibition of PKCα/β With Ruboxistaurin Antagonizes Heart Failure in Pigs After Myocardial Infarction Injury." Circulation Research **109**(12): 1396-1400.

Lara-Pezzi, E., J. Gómez-Salinero, A. Gatto and P. García-Pavía (2013). "The Alternative Heart: Impact of Alternative Splicing in Heart Disease." Journal of Cardiovascular Translational Research **6**(6): 945-955.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research **109**(12): 1332-1341.

Lin, Q., J. Schwarz, C. Bucana and E. N. Olson (1997). "Control of Mouse Cardiac Morphogenesis and Myogenesis by Transcription Factor MEF2C." Science **276**(5317): 1404-1407.

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

Linke, W. A. and S. Bucker (2012). "King of hearts: a splicing factor rules cardiac proteins." Nat Med **18**(5): 660-661.

Liu, Q., X. Chen, S. M. MacDonnell, E. G. Kranias, J. N. Lorenz, M. Leitges, S. R. Houser and J. D. Molkentin (2009). "Protein Kinase Cα, but Not PKCβ or PKCγ, Regulates Contractility and Heart Failure Susceptibility: Implications for Ruboxistaurin as a Novel Therapeutic Approach." Circulation Research **105**(2): 194-200.

Lu, G., S. Ren, P. Korge, J. Choi, Y. Dong, J. Weiss, C. Koehler, J.-n. Chen and Y. Wang (2007). "A novel mitochondrial matrix serine/threonine protein phosphatase regulates the mitochondria permeability transition pore and is essential for cellular survival and development." Genes & Development **21**(7): 784-796.

Lu, J., T. A. McKinsey, R. L. Nicol and E. N. Olson (2000). "Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases." Proceedings of the National Academy of Sciences **97**(8): 4070-4075.

Lu, J., T. A. McKinsey, C.-L. Zhang and E. N. Olson (2000). "Regulation of Skeletal Myogenesis by Association of the MEF2 Transcription Factor with Class II Histone Deacetylases." Molecular Cell **6**(2): 233-244.

Manetti, M., S. Guiducci, L. Ibba-Manneschi and M. Matucci-Cerinic (2011). "Impaired Angiogenesis in Systemic Sclerosis: The Emerging Role of the Antiangiogenic VEGF165b Splice Variant." Trends in Cardiovascular Medicine **21**(7): 204-210.

Martin, J. A. and Z. Wang (2011). "Next-generation transcriptome assembly." Nat Rev Genet **12**(10): 671-682.

McManus, C. J. and B. R. Graveley (2011). "RNA structure and the mechanisms of alternative splicing." Current Opinion in Genetics & Development **21**(4): 373-379.

Merdzhanova, G., S. Gout, M. Keramidas, V. Edmond, J. L. Coll, C. Brambilla, E. Brambilla, S. Gazzeri and B. Eymin (2010). "The transcription factor E2F1 and the SR protein SC35 control the ratio of pro-angiogenic versus antiangiogenic isoforms of vascular endothelial growth factor-A to inhibit neovascularization in vivo." Oncogene **29**(39): 5392-5403.

Mjaatvedt, C. H., H. Yamamura, A. A. Capehart, D. Turner and R. R. Markwald (1998). "TheCspg2Gene, Disrupted in thehdfMutant, Is Required for Right Cardiac Chamber and Endocardial Cushion Formation." Developmental Biology **202**(1): 56-66.

Molnár, A., A. Borbély, D. Czuriga, S. M. Ivetta, S. Szilágyi, Z. Hertelendi, E. T. Pásztor, Á. Balogh, Z. Galajda, T. Szerafin, K. Jaquet, Z. Papp, I. Édes and A. Tóth (2009). "Protein Kinase C Contributes to the Maintenance of Contractile Force in Human Ventricular Cardiomyocytes." Journal of Biological Chemistry **284**(2): 1031-1039.

Murphy, L. L., A. J. Moon-Grady, B. F. Cuneo, R. T. Wakai, S. Yu, J. D. Kunic, D. W. Benson and A. L. George Jr (2012). "Developmentally regulated SCN5A splice variant potentiates dysfunction of a novel mutation associated with severe fetal arrhythmia." Heart Rhythm **9**(4): 590-597.

Nalini Raghavachari, J. B., Yanqin Yang, Poching Liu, Kimberly Woodhouse, Daniel Levy, Chris O'Donnel, Peter J Munson and Gregory Kato (2012). "A systematic comparison and evaluation of high density exon arrays and RNA-seq technology used to unravel the peripheral blood transcriptome of sickle cell disease." BMC Medical Genomics **5**(28).

Olson, E. (2002). "A genetic blueprint for growth and development of the heart." Harvey Lect **98**: 41-64.

Olson, E. N. (2006). "Gene regulatory networks in the evolution and development of the heart." Science **313**(5795): 1922-1927.

Orenes-Piñero, E., S. Montoro-García, J. V. Patel, M. Valdés, F. Marín and G. Y. H. Lip (2013). "Role of microRNAs in cardiac remodelling: New insights and future perspectives." International Journal of Cardiology **167**(5): 1651-1659.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey and B. J. Blencowe (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nat Genet **40**(12): 1413-1415.

Qian, L., Y. Huang, C. I. Spencer, A. Foley, V. Vedantham, L. Liu, S. J. Conway, J.-d. Fu and D. Srivastava (2012). "In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes." Nature **advance online publication**.

Ramchatesingh, J., A. Zahler, K. Neugebauer, M. Roth and T. Cooper (1995). "A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer." Mol. Cell. Biol. **15**(9): 4898-4907.

Rana, M. S., V. M. Christoffels and A. F. M. Moorman (2013). "A molecular and genetic outline of cardiac morphogenesis." Acta Physiologica **207**(4): 588-615.

Refaat, M. M., S. A. Lubitz, S. Makino, Z. Islam, J. M. Frangiskakis, H. Mehdi, R. Gutmann, M. L. Zhang, H. L. Bloom, C. A. MacRae, S. C. Dudley, A. A. Shalaby, R. Weiss, D. M. McNamara, B. London and P. T. Ellinor (2012). "Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy." Heart rhythm : the official journal of the Heart Rhythm Society **9**(3): 390-396.

Sharma GD, N. H., Antonov AS, Gerrity RG, von Geldern T, Pandey KN (2002). "Expression of atrial natriuretic peptide receptor-A antagonizes the mitogen-activated protein kinases (Erk2 and P38MAPK) in cultured human vascular smooth muscle cells." Mol Cell Biochem **233**(2-3): 165-173.

Song, K., Y.-J. Nam, X. Luo, X. Qi, W. Tan, G. N. Huang, A. Acharya, C. L. Smith, M. D. Tallquist, E. G. Neilson, J. A. Hill, R. Bassel-Duby and E. N. Olson (2012). "Heart repair by reprogramming non-myocytes with cardiac transcription factors." Nature **485**(7400): 599-604.

Srivastava, D. (1997). "Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND." Nature Genet. **16**: 154-160.

Srivastava, D., P. Cserjesi and E. N. Olson (1995). "New subclass of bHLH proteins required for cardiac morphogenesis." Science **270**: 1995-1999.

Steinberg, S. F. (2008). "Structural Basis of Protein Kinase C Isoform Function." Physiological Reviews **88**(4): 1341-1378.

Sultan, M., M. H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf, M. Seifert, T. Borodina, A. Soldatov, D. Parkhomchuk, D. Schmidt, S. O'Keeffe, S. Haas, M. Vingron, H. Lehrach and M.-L. Yaspo (2008). "A Global View of Gene Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome." Science **321**(5891): 956-960.

Taegtmeyer, H., S. Sen and D. Vela (2010). "Return to the fetal gene program." Annals of the New York Academy of Sciences **1188**(1): 191-198.

Tischer, E., D. Gospodarowicz, R. Mitchell, M. Silva, J. Schilling, K. Lau, T. Crisp, J. C. Fiddes and J. A. Abraham (1989). "Vascular endothelial growth factor: A new member of the platelet-derived growth factor gene family." Biochemical and Biophysical Research Communications **165**(3): 1198-1206.

Turbendian, H. K., M. Gordillo, S.-Y. Tsai, J. Lu, G. Kang, T.-C. Liu, A. Tang, S. Liu, G. I. Fishman and T. Evans (2013). "GATA factors efficiently direct cardiac fate from embryonic stem cells." Development **140**(8): 1639-1644.

Ucar, A., S. K. Gupta, J. Fiedler, E. Erikci, M. Kardasinski, S. Batkai, S. Dangwal, R. Kumarswamy, C. Bang, A. Holzmann, J. Remke, M. Caprio, C. Jentzsch, S. Engelhardt, S. Geisendorf, C. Glas, T. G. Hofmann, M. Nessling, K. Richter, M. Schiffer, L. Carrier, L. C. Napp, J. Bauersachs, K. Chowdhury and T.

Thum (2012). "The miRNA-212/132 family regulates both cardiac hypertrophy and cardiomyocyte autophagy." Nat Commun **3**: 1078.

Wahbi, K., V. Algalarrondo, H. M. Bécane, V. Fressart, C. Beldjord, K. Azibi, A. Lazarus, N. Berber, H. Radvanyi-Hoffman, T. Stojkovic, A. Béhin, P. Laforêt, B. Eymard, S. Hatem and D. Duboc "Brugada syndrome and abnormal splicing of SCN5A in myotonic dystrophy type 1." Archives of Cardiovascular Diseases(0).

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth and C. B. Burge (2008). "Alternative isoform regulation in human tissue transcriptomes." Nature **456**(7221): 470-476.

Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nat Rev Genet **10**(1): 57-63.

Warf, M. B. and J. A. Berglund (2007). "MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T." RNA **13**(12): 2238-2251.

Wei Guo, S. J. B., Karla Esbona and Marion L. Greaser (2010). "Titin Diversity-Alternative Splicing Gone Wild." J Biomed Biotechnol.

Westerfield, M. (1995). "The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio)." Eugene: University of Oregon.

Witten, J. T. and J. Ule (2011). "Understanding splicing regulation through RNA splicing maps." Trends in Genetics **27**(3): 89-97.

Wu, Y., R. Dey, A. Han, N. Jayathilaka, M. Philips, J. Ye and L. Chen (2010). "Structure of the MADS-box/MEF2 Domain of MEF2A Bound to DNA and Its

Implication for Myocardin Recruitment." Journal of Molecular Biology **397**(2): 520-533.

Xiao, X., Z. Wang, M. Jang and C. B. Burge (2007). "Coevolutionary networks of splicing cis-regulatory elements." Proceedings of the National Academy of Sciences **104**(47): 18583-18588.

Xu, X., D. Yang, J.-H. Ding, W. Wang, P.-H. Chu, N. D. Dalton, H.-Y. Wang, J. R. Bermingham Jr, Z. Ye, F. Liu, M. G. Rosenfeld, J. L. Manley, J. Ross Jr, J. Chen, R.-P. Xiao, H. Cheng and X.-D. Fu (2005). "ASF/SF2-Regulated CaMKII[delta] Alternative Splicing Temporally Reprograms Excitation-Contraction Coupling in Cardiac Muscle." Cell **120**(1): 59-72.

Yae, T., K. Tsuchihashi, T. Ishimoto, T. Motohara, M. Yoshikawa, G. J. Yoshida, T. Wada, T. Masuko, K. Mogushi, H. Tanaka, T. Osawa, Y. Kanki, T. Minami, H. Aburatani, M. Ohmura, A. Kubo, M. Suematsu, K. Takahashi, H. Saya and O. Nagano (2012). "Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell." Nat Commun **3**: 883.

Yanazume, T., K. Hasegawa, T. Morimoto, T. Kawamura, H. Wada, A. Matsumori, Y. Kawase, M. Hirai and T. Kita (2003). "Cardiac p300 Is Involved in Myocyte Growth with Decompensated Heart Failure." Molecular and Cellular Biology **23**(10): 3593-3606.

Zhang, C. L., T. A. McKinsey, S. Chang, C. L. Antos, J. A. Hill and E. N. Olson (2002). "Class II Histone Deacetylases Act as Signal-Responsive Repressors of Cardiac Hypertrophy." Cell **110**(4): 479-488.

Zhang, S.-S. and R. M. Shaw (2013). "Multilayered regulation of cardiac ion channels." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1833**(4): 876-885.

FIGURE LEGENDS:

Figure 1: Common Pattern of Alternative RNA Splicing Associated with Cardiac Development and Heart Failure**.**
(A) Relative expression levels of 30 exons were measured by real-time RT-PCR using exon specific primers. The RNA samples were obtained from left ventricles of mouse P1 neonatal hearts (Neonatal), 5 month old normal adult hearts (Adult) and age-matched pressure-overload induced failing hearts (Failing Heart) as indicated. The list of individual exons and the expression levels were illustrated in Supplemental Figure 2(n=3 for each group). (B) Schematic view of Mef2a gene structure, alternatively spliced transcripts including exon 5a or 5b, and Mef2a protein with functional domains illustrated. (C)Relative expression ratio of Exon a over Exon b in Mef2a, Mef2c and Mef2d genes in P1 mouse neonatal heart, 5 month mouse adult heart and age matched pressure-overload induced failing hearts as indicated. *, $p<0.05$.(D) Relative expression ratio of Exon a over Exon b in Mef2a, Mef2c and Mef2d genes in human fetal hearts (n=4), non-failing adult hearts (Adult, n=4) and dilated cardiomyopathy hearts (n=4). *, $p<0.05$.

Figure 2: **RBFox1 is dynamically regulated in Cardiac Development and Heart Failure.**

(A) qRT-PCR for RBFox1 mRNA levels normalized to GAPDH in Sham operated and 8 weeks post-TAC induced mouse hearts (n=3 each). *, $p<0.05$.(B) Immunoblot for RBFox1 protein levels in Sham operated and 8 weeks-post TAC induced mouse hearts. GAPDH was used as loading control. (C) qRT-PCR for RBFox1 mRNA in P1 neonatal mouse hearts and 5 monthsold adult mouse hearts. **, $p<0.01$.(D) Immunoblot for RBFox1 protein levels in P1 neonatal mouse hearts and 5 months old adult mouse hearts. (E) qRT-PCR for human RBFox1 mRNA in fetal heart and adult non-failing hearts, (n=4 for each group), **, $p<0.01$. (F) qRT-PCR for zRBFox1 mRNA in zebrafish embryos at 48 hours, 72 hours post fertilization (hpf) and 2 months old adult fish. *, $p<0.05$.

Figure 3: RBFox1 is Essential for Cardiac Development and Function
(A-C) Representative images of *Cmlc:GFP* zebrafish at 48 hour post-fertilization, injected with vehicle (A), RBFox1-MO alone (B) or in combination with zebrafish RBFox1 mRNA (C). Inserts are high magnification images showing pericardial edema and collapsed heart tube in RBFox1-MO fish. (D-F) Representative images of the same zebrafish hearts visualized under fluorescent microscope.(G-I) The M-mode tracing of zebrafish ventricle obtained from recorded video images (see Supplemental Information) using LQ program described earlier(Lu, Ren et al. 2007). (J) Quantification of zebrafish cardiac function measured as ejection fraction from Control, RBFox1 morphants and RBFox1 morphants plus RBFox1 mRNA injected embryos. **, $p<0.01$. (K) Average heart rates of 28hpf Control zebrafish, RBFox1-morphants and RBFox1-morphants plus RBFox1 mRNA injected embryos, **, $p<0.01$. (L) Summary of the number of embryos received morpholino injection. The number of embryos developed abnormal pericardial edema was recorded. (M) Representative images of cardiac gene expression pattern in zebrafish illustrated by whole-mount in situ hybridization for Notch1b, amhc and vmhc in Control and RBFox1 morphants as indicated. The red arrows indicate atrial-ventricular septum, and the red dashed lines indicate ventricular and atrial junction.

Figure 4: **RBFox1 is a Sufficient and Necessary Trans-Acting Regulator for MEF2 Splicing in Heart.**

(A) Schematic view of conserved RBFox1 binding motif in MEF2 pre-mRNA sequences from mouse, human, zebrafish and fugu fish adjacent to the alternatively spliced exons. Note: During evolution, zebrafish genome contains two separate Mef2c genes, Mef2c-a and Mef2c-b. (B) Relative expression ratio between Exon a and Exon b of zMef2a and zMef2d in zebrafish hearts at 48hpf and 72hpf and 2 months old adult zebrafish hearts. * $p<0.05$. (C) RBFox1 mRNA is efficiently overexpressed in neonatal rat ventricle cardiomyocytes (NRVM) using adv-RBFox1 vector as measured by qRT-PCR. (D) Relative expression ratio between exon-5a (E5) vs. exon-5b of Mef-2a in control and RBFox1 expressing NRVM,* $p<0.05$. (E) Relative expression ratio between exon-4a (E4) vs. exon -4b of Mef-2d in control and RBFox1 expressing NRVM,* $p<0.05$. (F) Relative expression ratio between exon-a vs. exon-b of Mef-2a and Mef2d in Control and RBFox1-Morphant zebrafish hearts, * $p<0.05$. (G) RBFox1 directly regulates MEF2 alternative splicing. Minigene reporter constructs containing mouse Mef2a Exon 5b and adjacent intron fragment containing putative RBFox1 binding motif (UGCAUG), mutant RBFox1 binding motif (UUCGUA) and deleted RBFox1 binding motif are illustrated. Reporter construct was transfected into HEK293 cells alone or in combination with RBFox1 expressing vector. 48 hours post transfection, the treated cells were harvested and semi-quantitative RT-PCR was performed to determine the relative level of the transcripts containing exon E5b or excluded exon E5b. GAPDH mRNA level was used as control. RBFox1 expression was confirmed by immunoblot with HDAC2 used as internal loading control.

Figure 5: Differential Effects of Mef2a Splicing Variants in Zebrafish Development

(A) Representative image of zebrafish embryos injected with control morpholino showing normal morphology at 48hpf. (B) Representative image of 48hpf zebrafish injected with different dosages of zebrafish Mef2a-4a mRNA transcript as indicated. (C) Representative image of 48hpf zebrafish injected with different dosages of zebrafish Mef2a-4b mRNA transcript as indicated. (D) Representative image of 48hpf zebrafish injected with different dosages of mouse Mef2a-5a mRNA transcript as indicated. (E) Representative image of 48hpf zebrafish injected with different dosages of mouse Mef2a-5b mRNA transcript as indicated.

Figure 6: **Functional Role of MEF2 Splicing Variants in RBFox1 Regulated Cardiac Development**.
(A) Representative images of zebrafish at 48hpf injected with Control, RBFox1, zMef2a-4a, zMef2a-4b morpholinos alone or in combination as indicated. Enlarged inserts shows pericardial edema in zRBFox1 morphants and zRBFox1 plus zMef2a-E4b morphants. (B) Summary data showing the number of embryos received each type of injection and the number of embryos developed pericardial edema. (C) Representative images of cardiac gene expression pattern in zebrafish illustrated by whole-mount in situ hybridization for Notch1b, amhc and vmhc in Control, RBFox1 and RBFox1+zMef2a-E4a morphants as indicated. The red arrows indicate atrial-ventricular septum, and the red dashed lines indicate ventricular and atrial junction.

Figure 7:**Mef2a Splicing Variants Have Distinct Transcriptional Activity**

(A) Schematic view of the experiment design. Zebrafish embryos were injected with individual zebrafish Mef2a splicing variants containing Exon4a or Exon4b. The embryos were collected at 24hpf and RNA was extracted for RNA-Seq. (B)Gene activation profile

in zebrafish embryos overexpressing zMef2a splicing variant containing Exon4a or Exon4b. The number on top indicates total number of genes in each group representing genes induced or suppressed by both or either splicing variants as indicated. (C)Phase-contrast light microscopic images of NRVM expressing mouse Mef2a-Exon 5a or Exon5b. (D)Relative mRNA levels of Nkx2.5, Gata4, Tbx20 and total Mef2d in NRVM 48hr post infection of adenoviral vectors expressing mouse Mef2a-Exon5a or Mef-2a-Exon5b (n=3 for each sample). * $p<0.05$. (E) Transcriptional activities of Nkx2.5 promoter constructs with different lengths as illustrated were measured based on luciferase activities in NRVM. The locations of the putative MEF2 binding motifs are identified at -389 and -169 bp as indicated.* $p<0.05$ vs. Control NRVM, # $p<0.05$ between Mef2a-E5a and Mef2a-E5b.

Figure 8:**RBFox1 Mediated RNA Splicing in Cardiac Hypertrophy and Heart Failure**

(A) Cellular morphology of NRVM expression RBFox1 48 hours post adenoviral infection. (B) β–Myosin heavy chain (MHC) and atrial natriuretic factor (ANF) expression in NRVM expressing RBFox1 at 48 hours post infection. (C) qRT-PCR for human RBFox1 mRNA in fetal heart and adult non-failing hearts, (n=4 for each group), **, $p<0.01$.(F)qRT-PCR for human RBFox1 mRNA in non-failing adult hearts and dilated cardiomyopathy hearts (n=4 for each group), *, $p<0.05$. (D) qRT-PCR for RBFox1 expression in Sham operated, 14day post-TAC, 14day TAC followed by 1-day dTAC (Experimental Procedure for Details). (E) RBFox1 protein expression in the left ventricles of transgenic mouse hearts (RBFox1-TG) before and after TAC surgery as detailed in Experimental Procedures. β-catenin was used as internal loading control. (F) Left ventricle (LV) weight/body weight ratio in RBFox1 Transgenic mice and non-transgenic littermates before and after 6-weeks of TAC. *, $p<0.05$. (G) Representative M-mode echocardiogram of a RBFox1 transgenic heart and non-transgenic littermate before and after TAC. (H) Cardiac ejection fraction of RBFox1-TG and non-TG control mice before and after different time period post-TAC as indicated. *, $p<0.05$. (I) ANF and β-MHC expression in RBFox1 transgenic and non-transgenic littermate control hearts before and after TAC. *, $p<0.05$

Table 1:Enrichment of RNA splicing factor binding motifs for alternatively spliced exons in cardiac transcriptome

Summary of enriched binding motifs among exons differentially included normal or failing heart transcriptome identified through RNA-seq (Lee, Gao et al. 2011). Using de novo motif discovery, 5-mers that are both evolutionarilyconserved and highly enriched among the exons in the flanking introns and exonic regions. Five regions for each exon are analyzed and annotated: upstream intron first 250 nt (UpIn 1st), upstream intron last 250nt (UpIn 2nd), downstream intron first 250nt (DnIn 1st), downstream intron last 250nt (DnIn 2nd). RBFox binding motifs are highlighted in red.

Figure 5 Graphic abstract

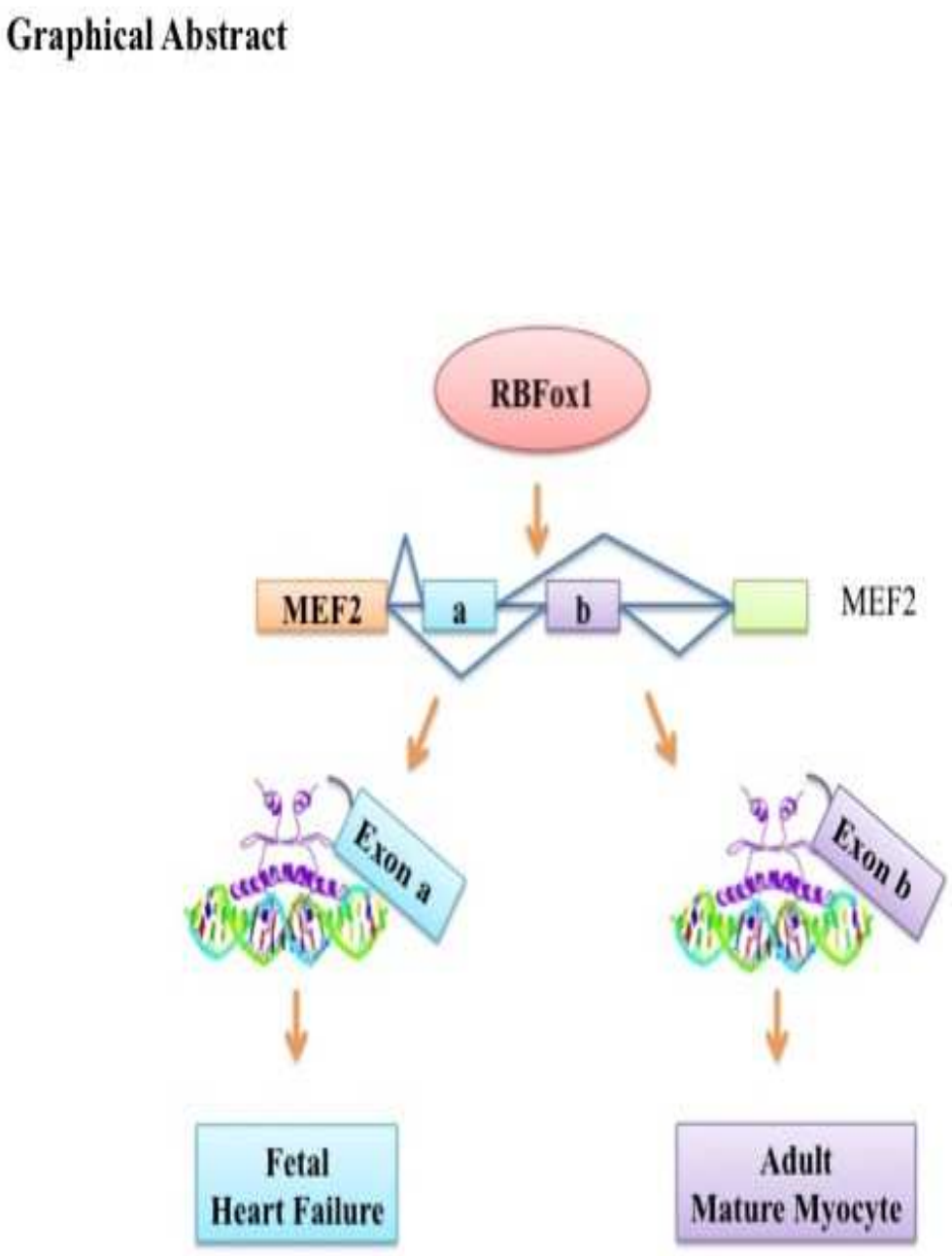
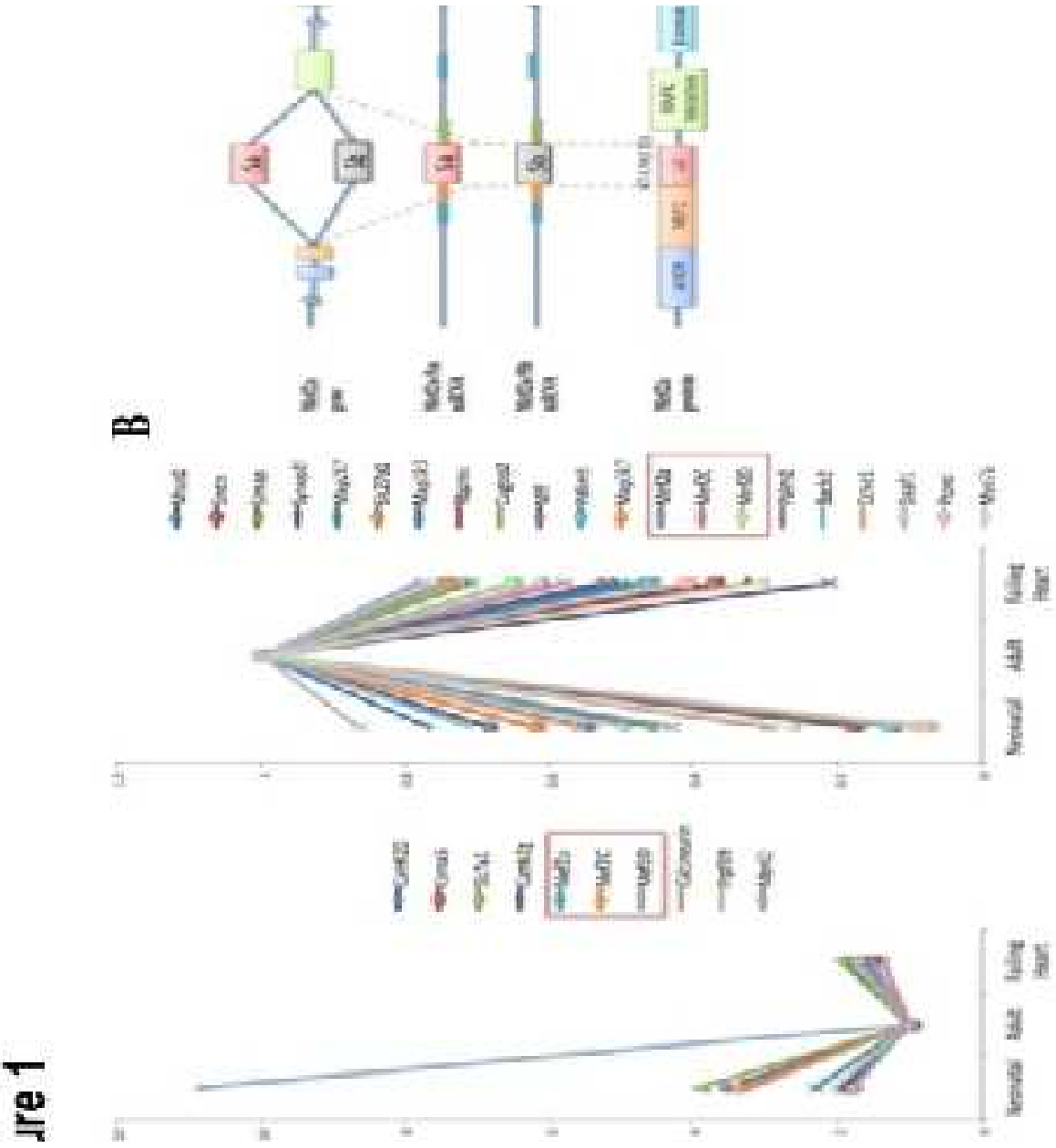


Figure 5.1 Fetal like alternative splicing associated with heart disease



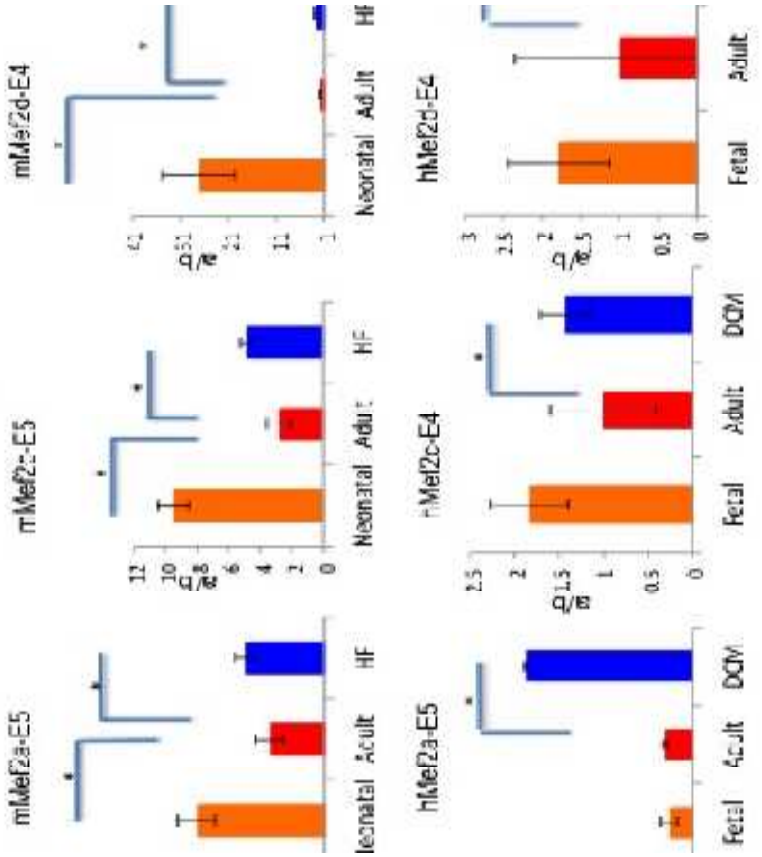


Figure 5.2 RBFox1 expression is dynamically regulated during cardiac development and disease remodeling

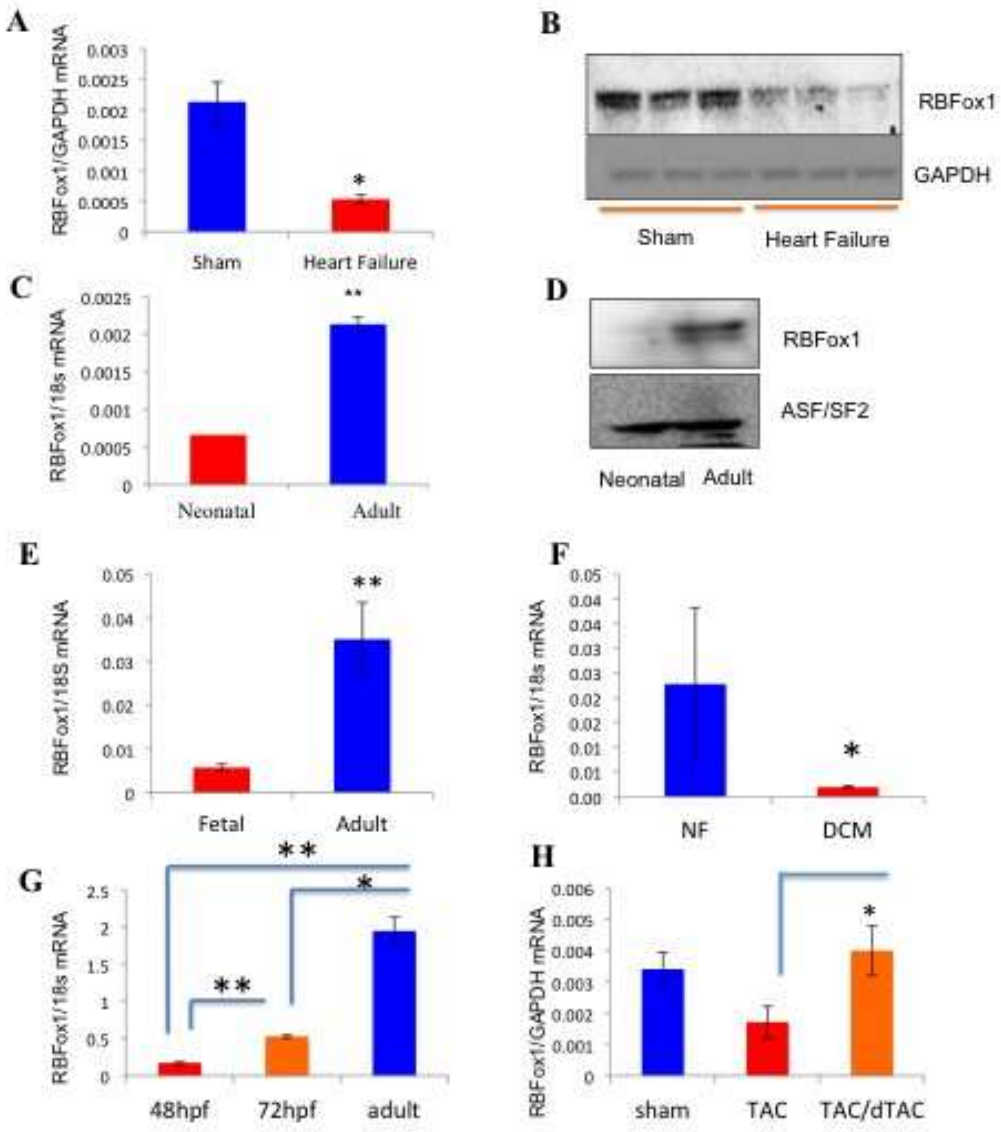


Figure 5.3 RBFox1 is critical for zebrafish cardiac function

Figure 3

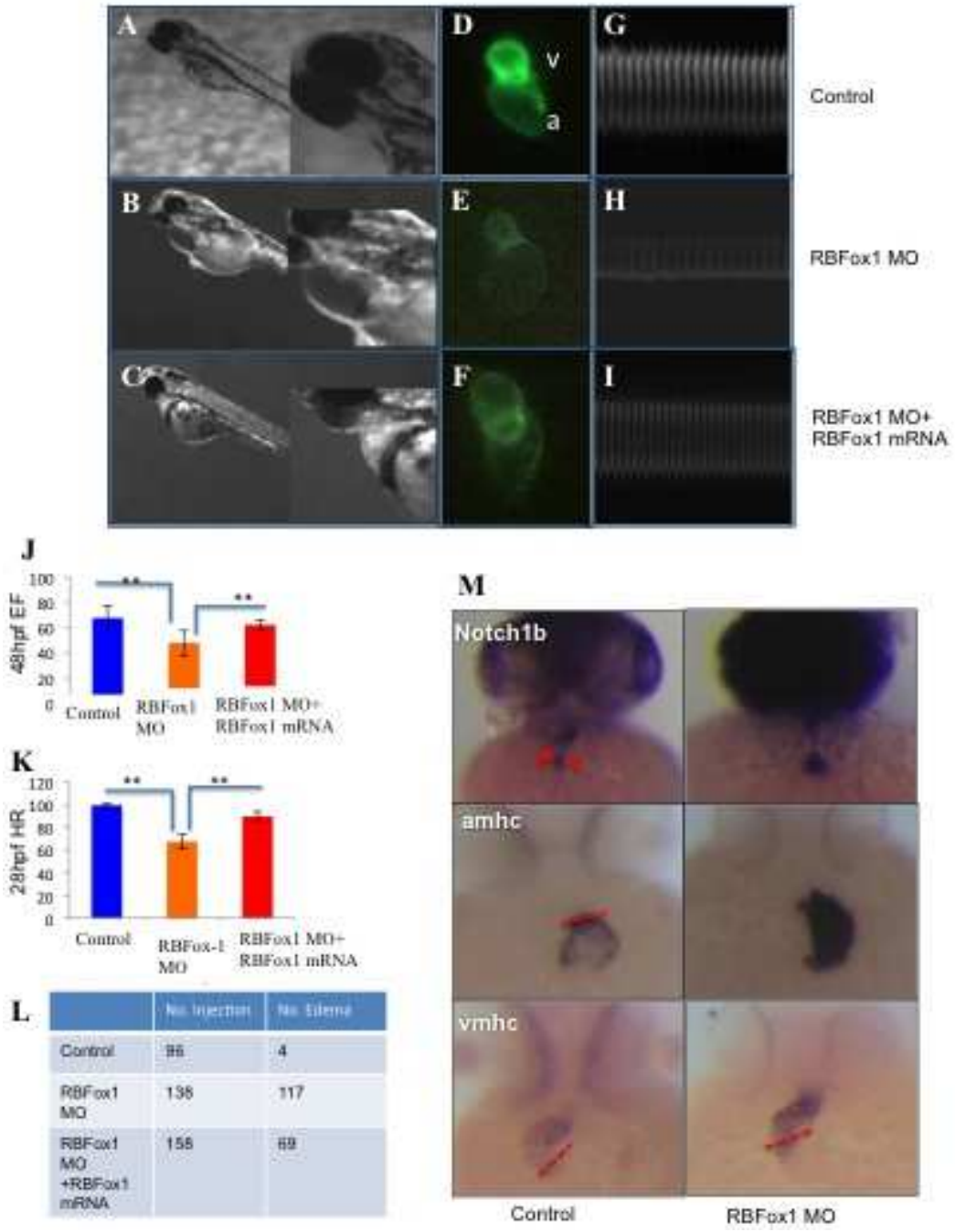


Figure 5.4 MEF splicing is directly regulated by RBFox1

Figure 4

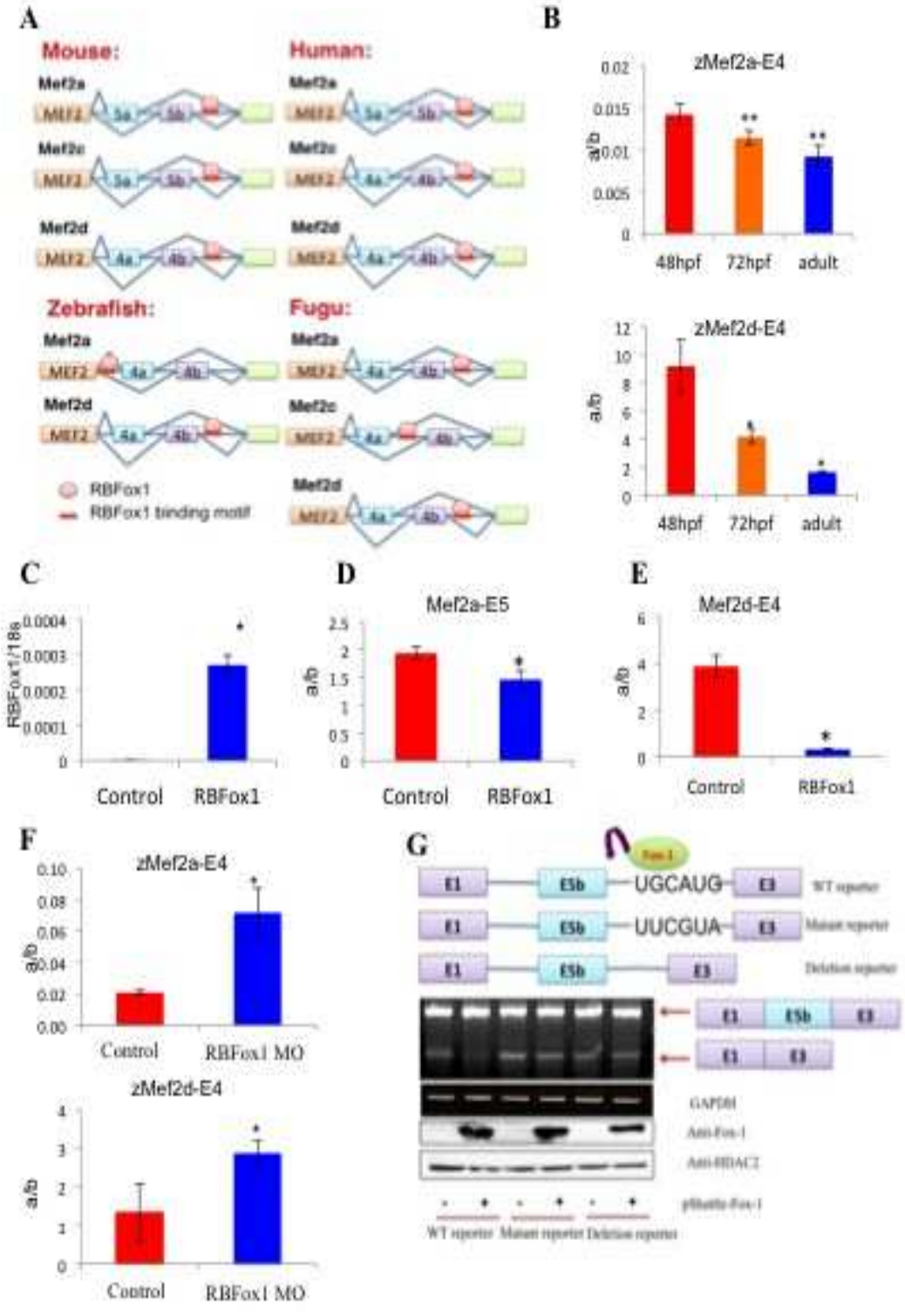

A
Mouse:
Human:
Zebrafish:
Fugu:
Mef2a
Mef2c
Mef2d
MEF2
RBFox1
RBFox1 binding motif
B
zMef2a-E4
zMef2d-E4
a/b
48hpf
72hpf
adult
C
RBFox1/18s
Control
RBFox1
D
Mef2a-E5
E
Mef2d-E4
F
zMef2a-E4
zMef2d-E4
Control
RBFox1 MO
G
UGCAUG
UUCGUA
WT reporter
Mutant reporter
Deletion reporter
GAPDH
Anti-Fox-1
Anti-HDAC2

Figure 5.5 MEF2 splicing variants have differential functional impact on zebrafish development

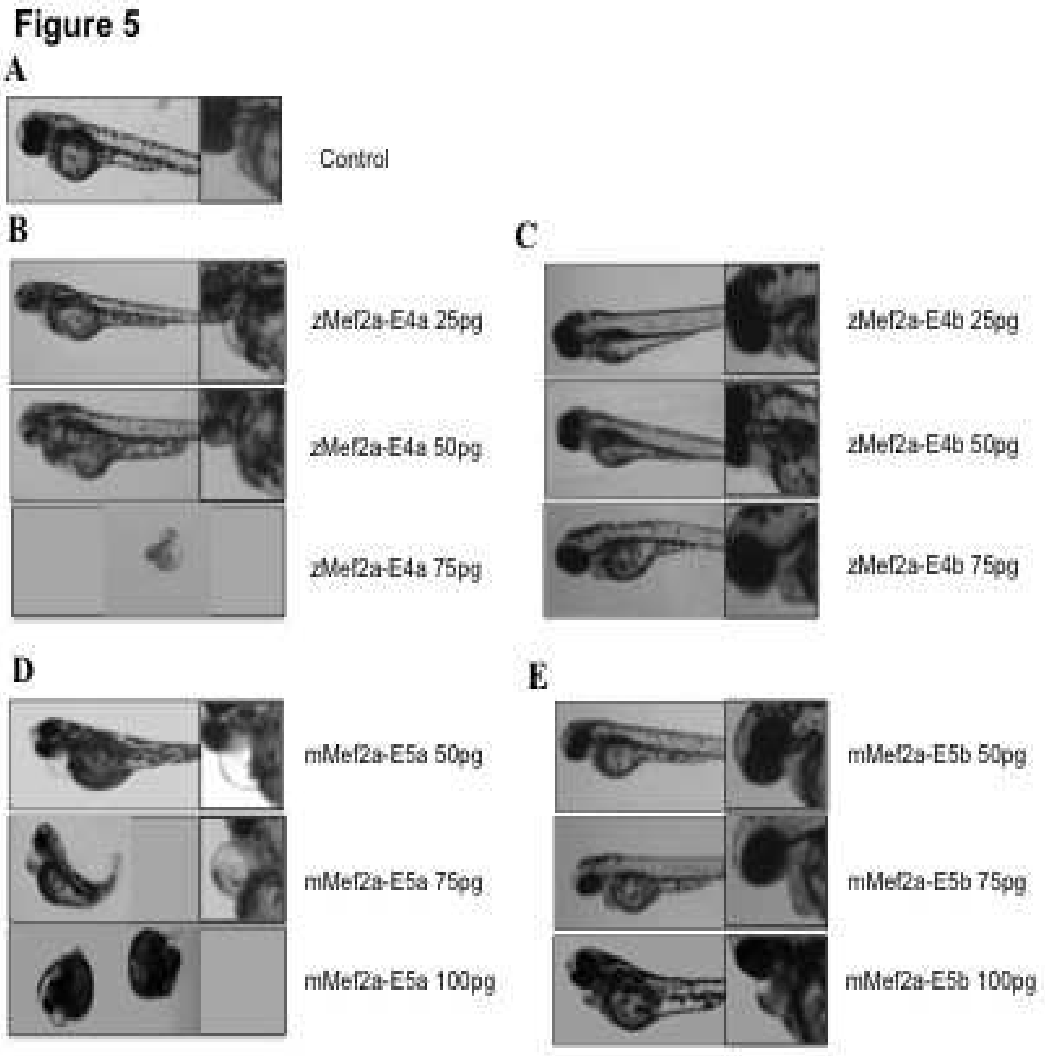
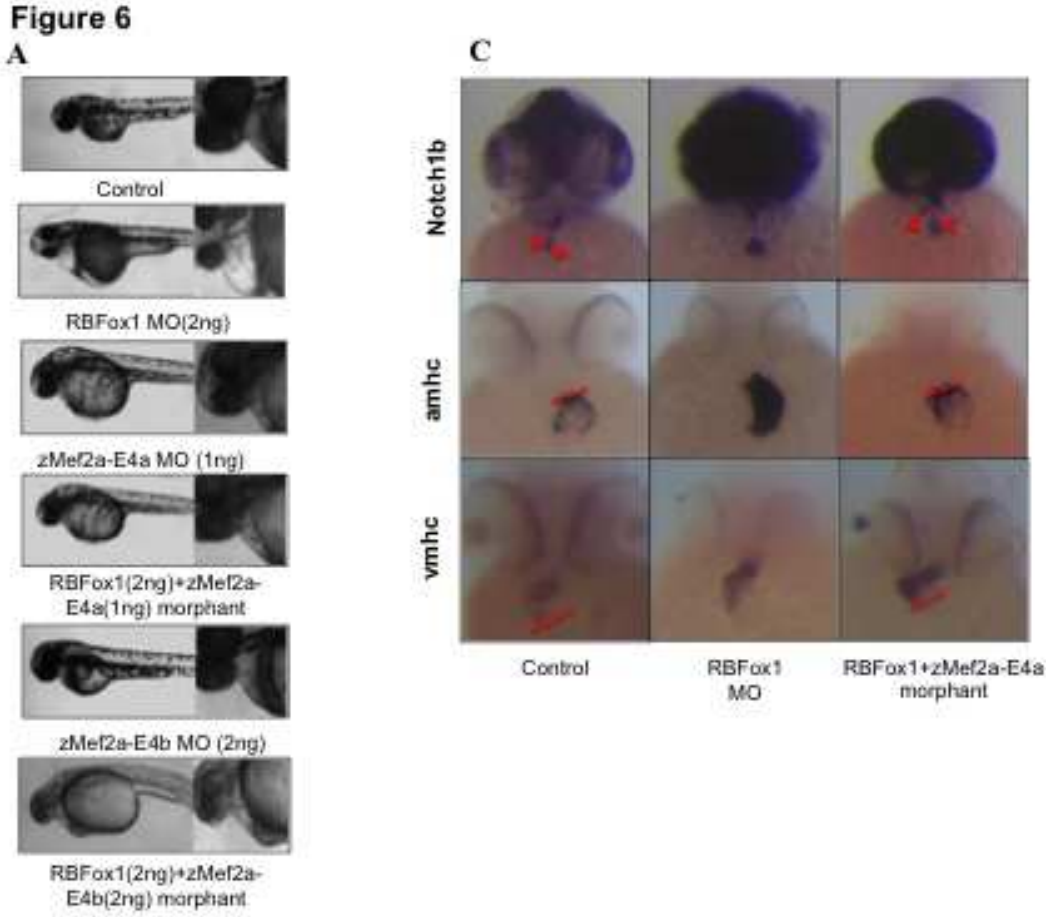


Figure 5.6 MEF2 splicing is critical for RBFox1 mediated splicing regulation



B

Sample	No. Injection	No. Edema
Control	300	7
RBFox1-MO	400	353
RBFox1-MO+Mef2a-E4a MO	186	97
RBFox1-MO+Mef2a-E4b MO	132	119

Figure 5.7 MEF2 splicing variants have different transcription activities

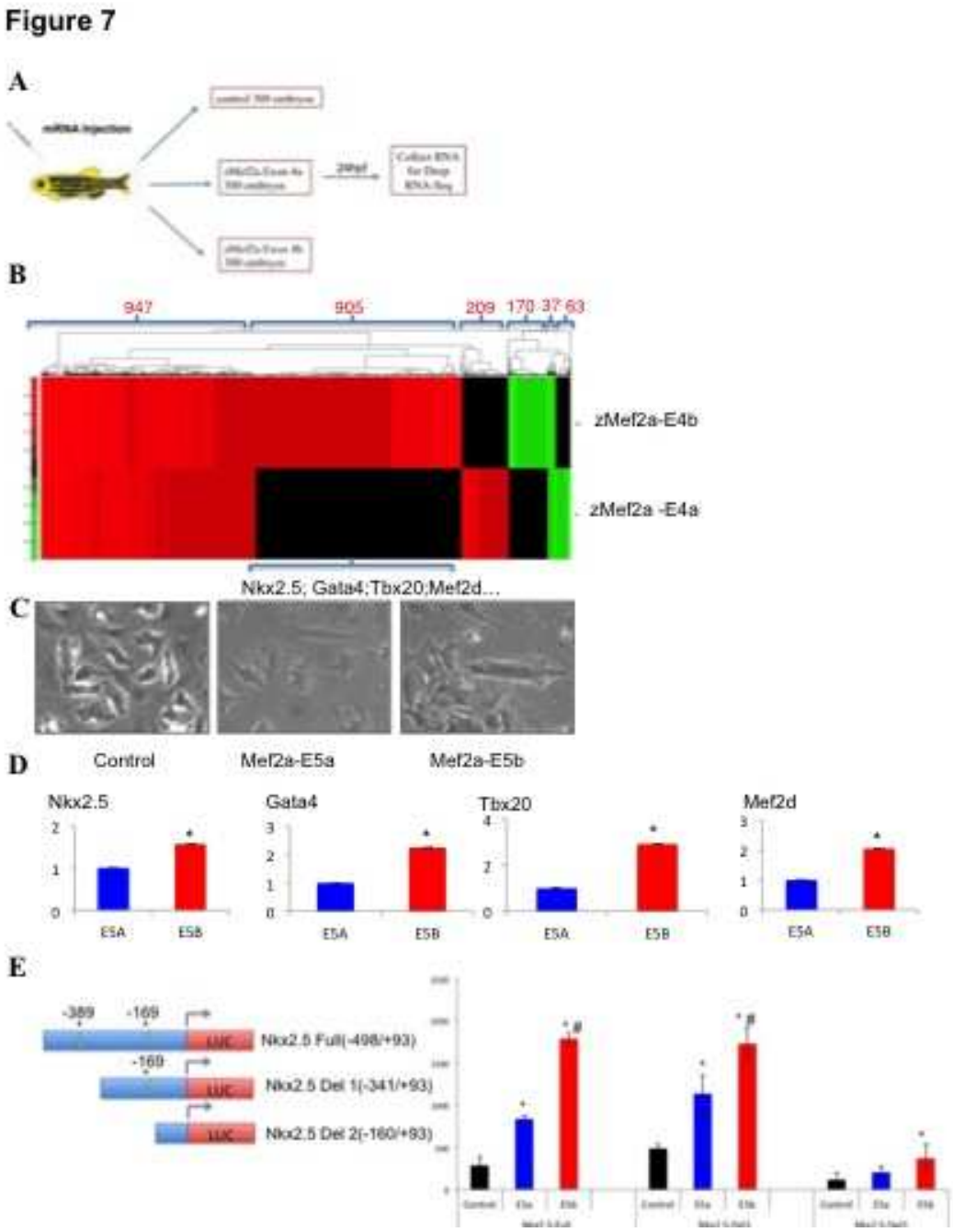


Figure 5.8 Restoring RBFox1 expression is sufficient to preserve cardiac function in pressure-overload

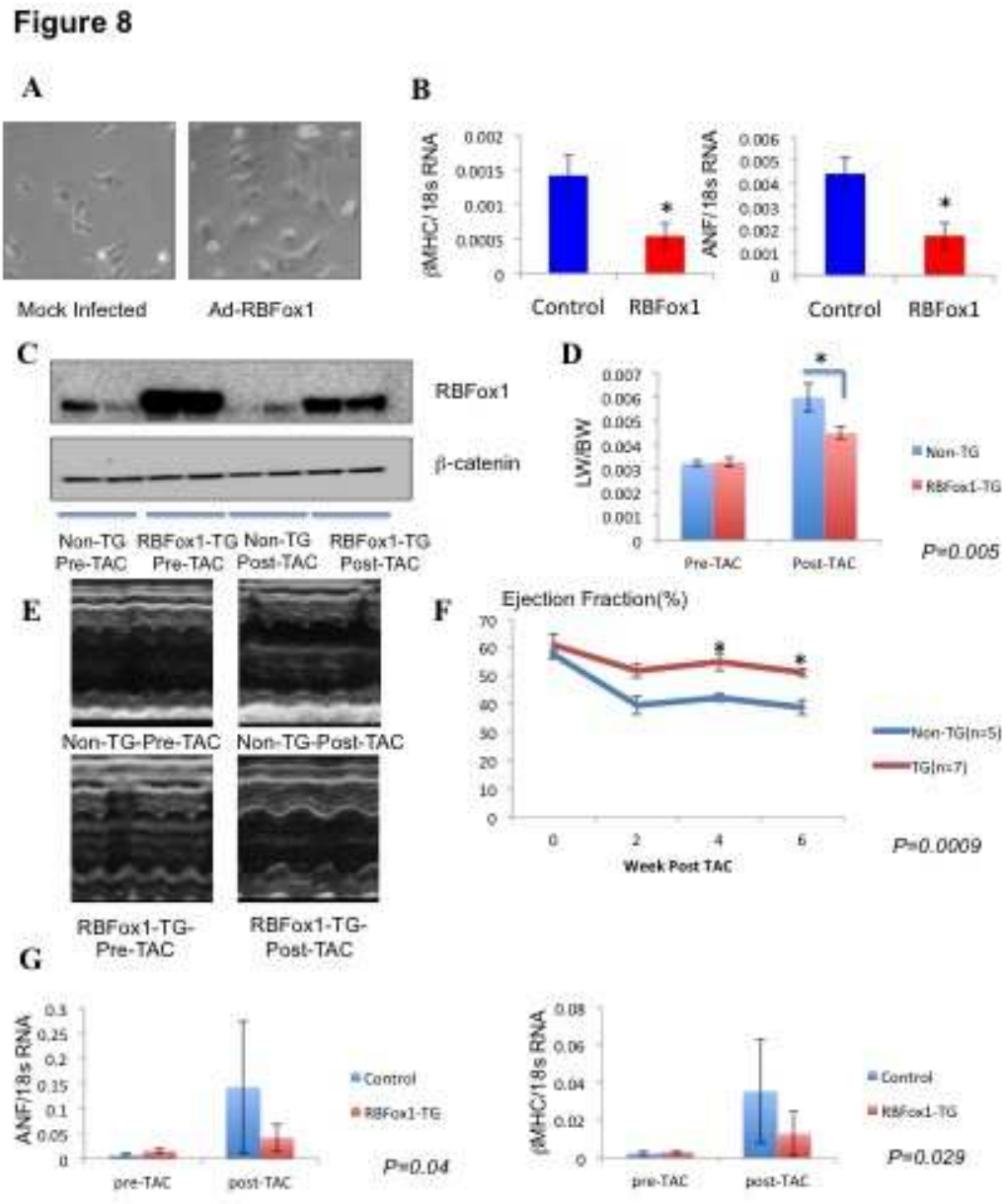


Table 5.1 Enriched motif in alternative splicing in failing hearts

Table 1

Normal

Upln 1st	Kmer	Conservation	Enrichment	Known?
	CGGCG	7.21E-005	6.14E-004	
	GCGGC	1.90E-003	2.31E-002	
	CCGCC	2.74E-003	1.34E-002	
Upln 2nd				
	TTTTT	1.23E-004	8.78E-014	FESS:TIA
	TGATA	2.30E-003	8.58E-003	
	CACGT	6.01E-003	1.09E-002	
Exon				
	TACCA	1.75E-005	2.07E-004	
	TGTGT	3.32E-004	8.79E-003	CELF;
	AAGGA	6.39E-004	6.46E-004	RESE;
Dnln 1st				
	TGCAT	3.26E-006	3.33E-005	RBFOX;
	GCATG	1.32E-007	1.36E-006	RBFOX;
	CTGCT	1.48E-004	1.94E-002	
Dnln 2nd				
	ACATA	5.60E-004	9.89E-004	
	TGTGT	7.49E-004	3.93E-005	CELF;
	CGCGG	7.68E-004	6.12E-003	

Heart Failure

Upln 1st	Kmer	Conservation	Enrichment	Known?
	CACAC	9.25E-004	7.74E-005	hnRNP L?
	CGTGG	5.89E-003	1.25E-003	
	GAGCA	8.03E-003	1.14E-002	
Upln 2nd				
	CTAAC	2.62E-004	9.62E-003	SF1;
	ACGCA	2.27E-003	3.00E-002	RESE;
	TATTT	2.84E-003	5.03E-003	
Exon				
	AAGGA	8.95E-007	6.94E-004	RESE;
	AGAAG	9.22E-007	1.26E-004	RESE;
	CTGGA	4.10E-006	7.57E-005	RESE;
Dnln 1st				
	CGCAC	2.67E-006	7.85E-004	
	GCGCA	5.02E-005	9.42E-003	
	CAGAA	5.39E-004	1.08E-003	RESE;
Dnln 2nd				
	GCATG	5.48E-005	1.70E-002	RBFOX;
	AAACA	5.01E-004	6.81E-006	RESE;
	TTTAA	4.91E-003	9.38E-003	

Table 5.1.S List of verified alternative splicing in failing and neonatal hearts

Supplement Table 1

Alternative Splicing Events Candidate Genes

Exon Inclusion		Exon Exclusion
CamK2D		Mtus1
Cyma5		Smtn
Slc7a1		Slmap
Camk2g		Synop2l
Mef2a		Map2k7
Mef2C		Slc27a1
Mef2D		Map3k3
Calcineurin		Rbpms
Prpf39		Cugbp2
Mbnl2		Mff
		Pdlim5
		Map3k7
		Mef2a
		Mef2C
		Mef2D
		Palm2
		Bach2
		Ccne1
		Deaf1
		Ptprd
		Myo7a
		Lyst

Figure 5.1.S Cardiac function analysis for mouse hearts used for deep RNA-sequencing

Figure S1

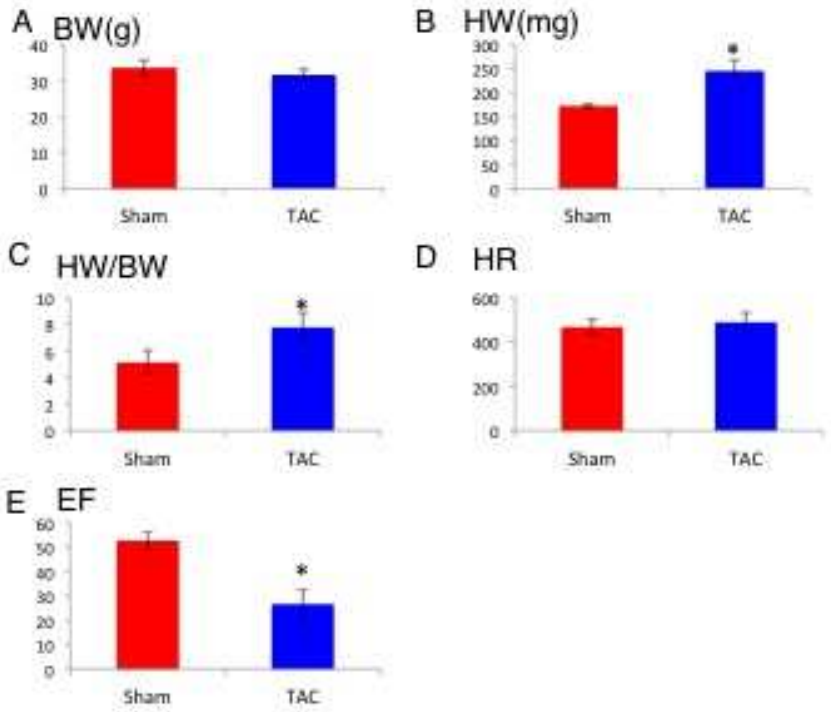


Figure 5.2.S Individual alternative splicing event confirmed in neonatal and diseased heart

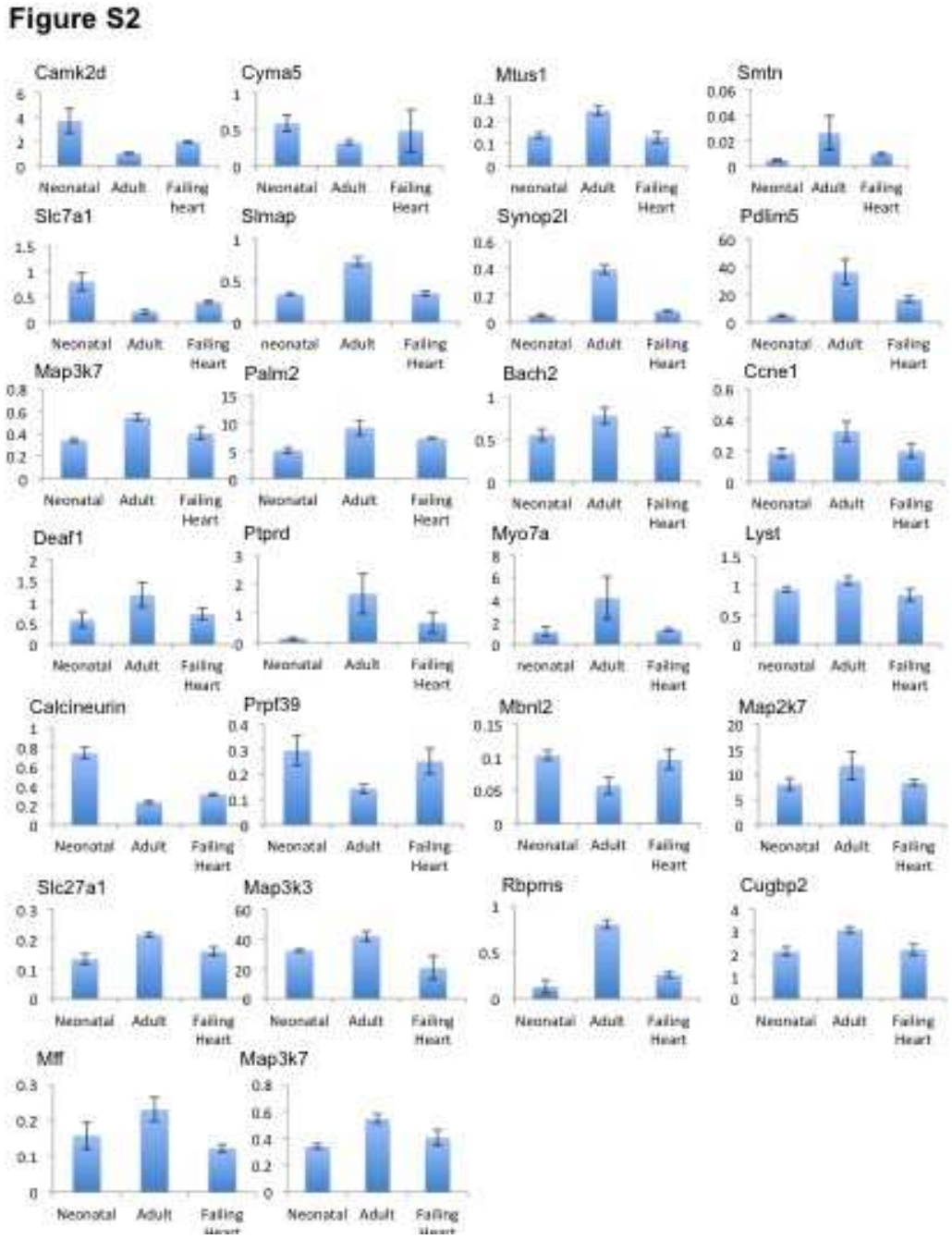


Figure 5.3.S Expression profile of alternative splicing regulators

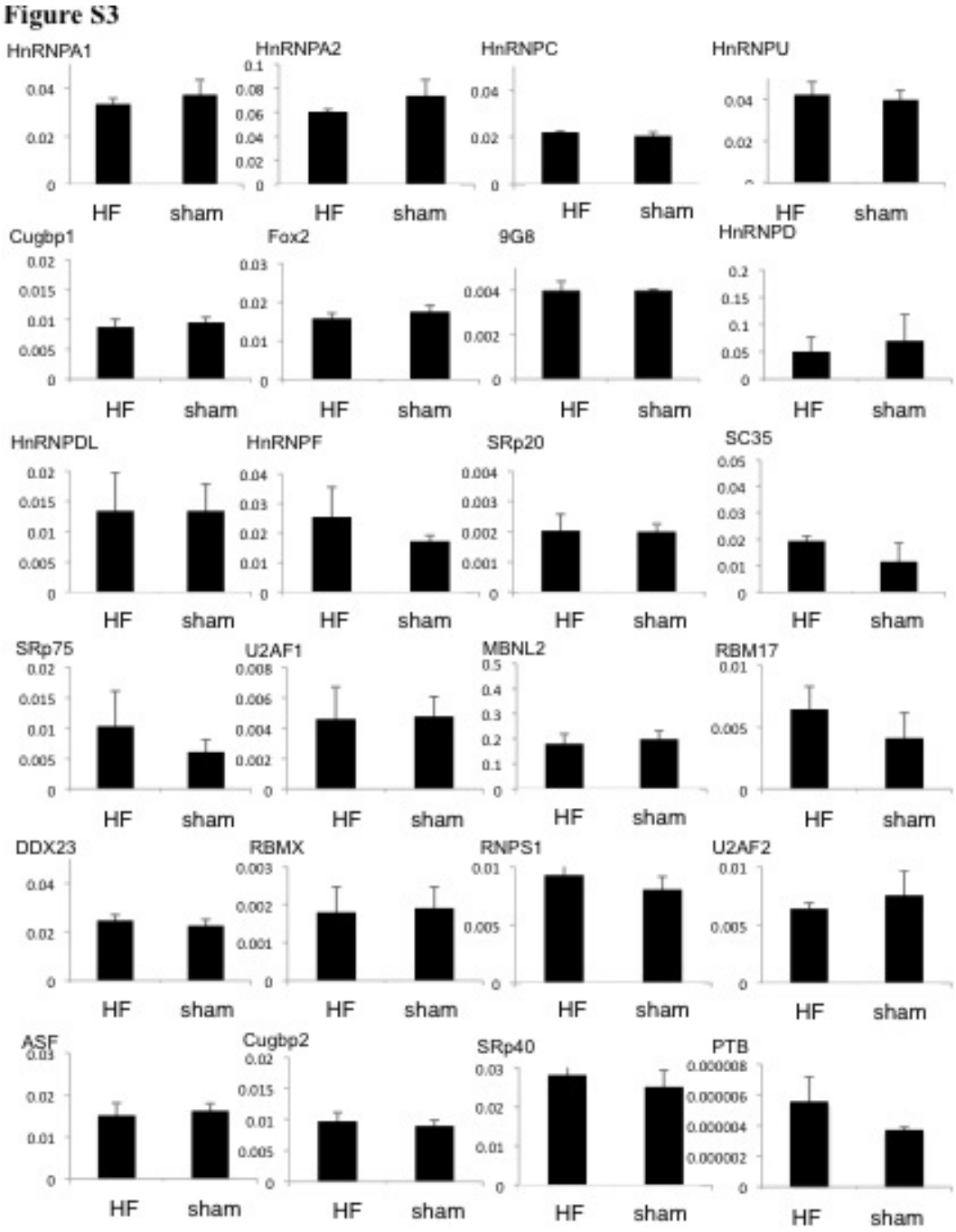


Figure 5.4.S Hypertrophy marker gene expression of mouse hearts during TAC and dTAC experiment

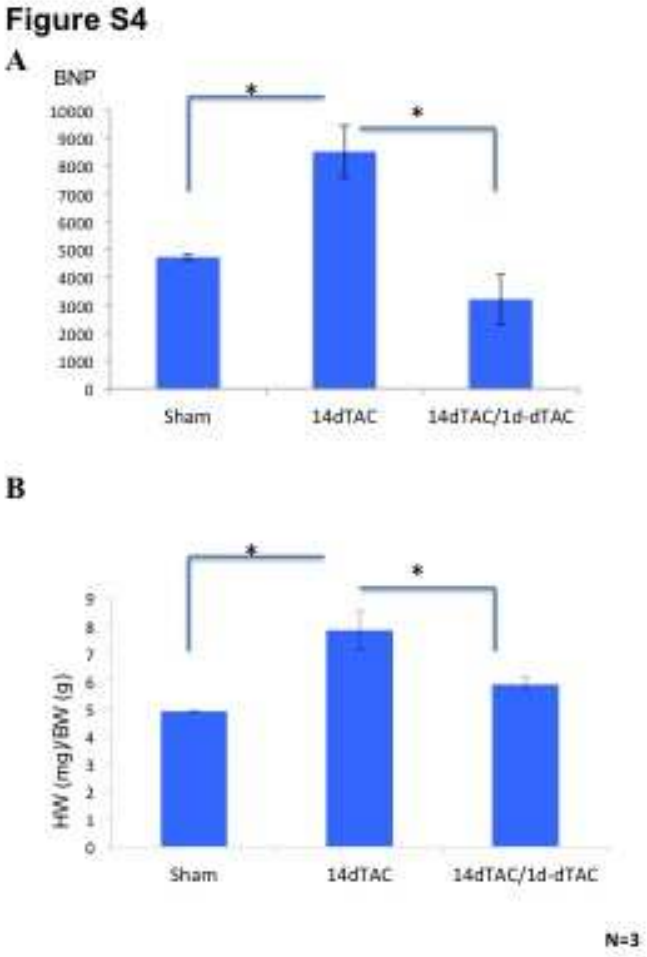


Figure 5.5.S Experimental design for zebrafish cardiac development gene expression analysis

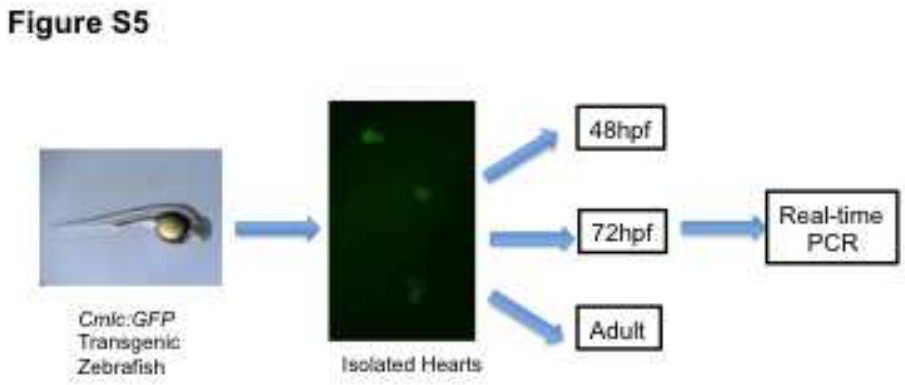


Figure 5.6.S RBFox1 morphant phenotype can be rescued by mouse RBFox1 mRNA

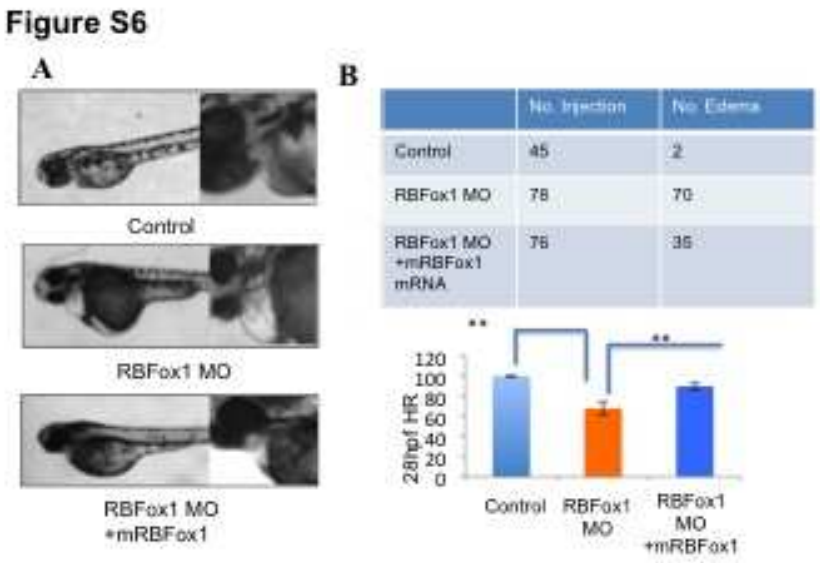


Figure 5.7.S Minigene reporter analysis of Mef2d

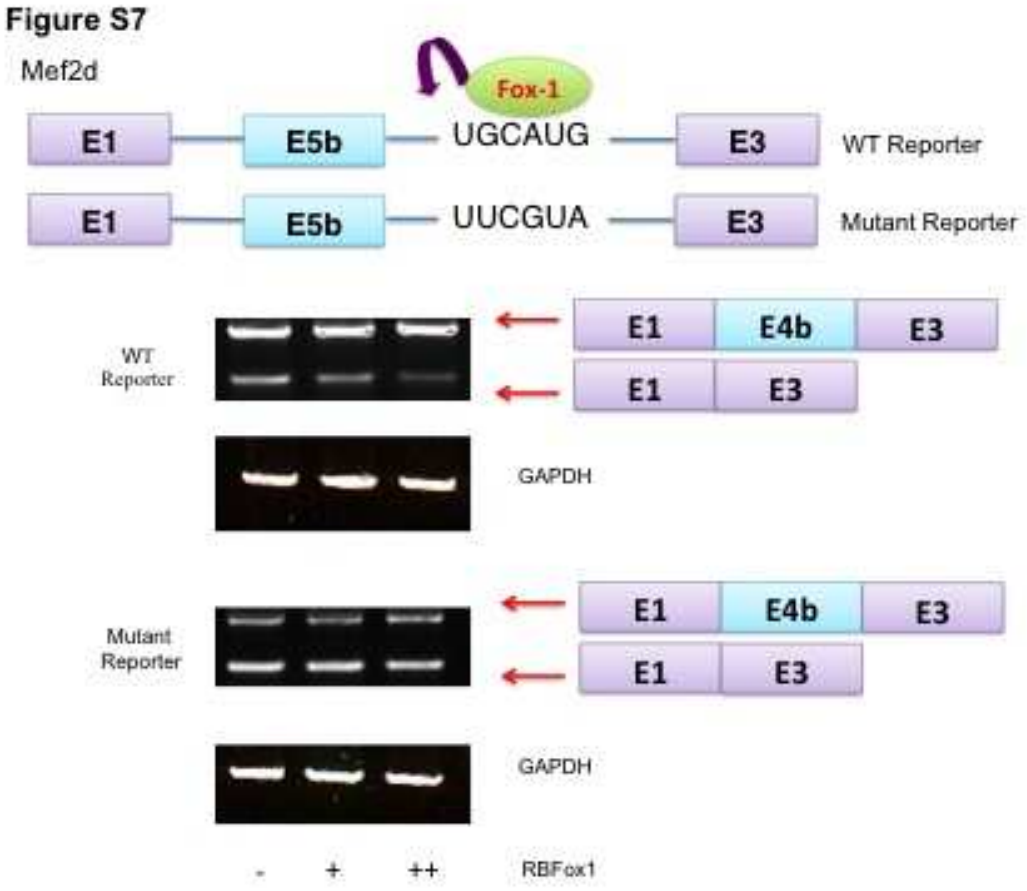


Figure 5.8.S Mef2a splicing variants can be specifically knocked down in zebrafish

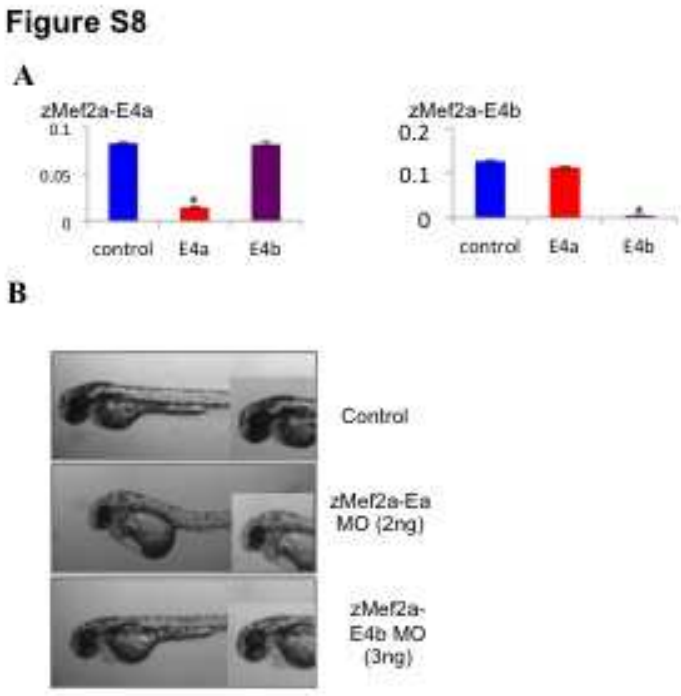


Figure 5.9.S Mef2a splicing variants have different activities on Tbx20 Luciferease reporter

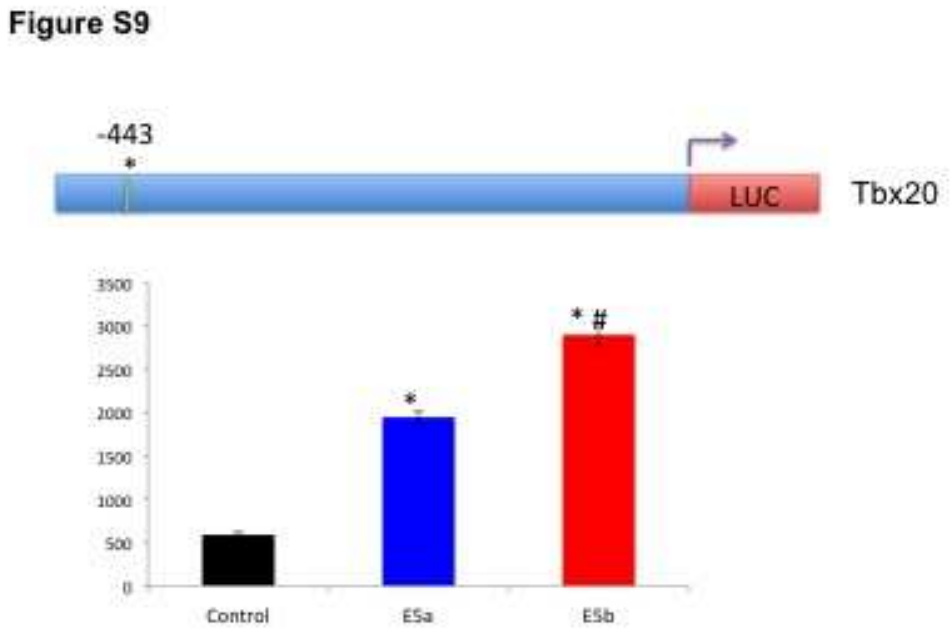
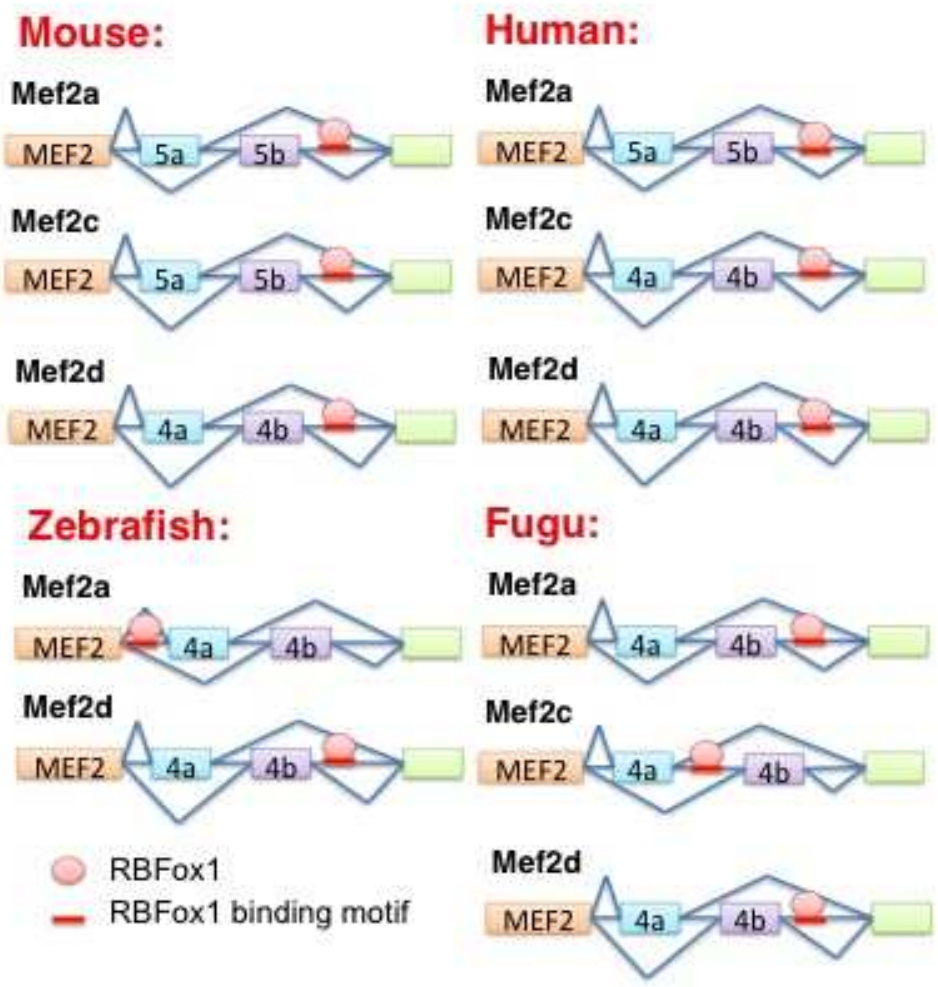


Figure 5. 10.S Schematic view of MEF2 pre-mRNA sequence



Chapter Six

Discussion and Future Direction

Summary

Our study has provided a comprehensive understanding of cardiac transcriptome during maturation and pathological remodeling. Based on our deep RNA-sequencing analysis, we have identified previously un-annotated splicing events, lncRNAs and novel transcript clusters. Using PKCα-NE as example, we have further demonstrated these novel splicing variants could have functional impact during cardiac pathological remodeling. Lastly, we have identified a critical splicing regulator—RBFox1 that is dynamically regulated during cardiac transcriptome maturation and disease remodeling. By establishing a link between RBFox1 and cardiac transcription factor—MEF2, we have identified a novel cardiac regulatory circuit at alternative splicing level.

Our study has several folds of significances that will be listed in this chapter.

A comprehensive understanding of the cardiac transcriptome

During the past decades, with microarray studies, we have generated detailed blueprint based on transcription factors during cardiac development, including key transcription factors: MEF2, Nkx2.5, GATA and TBX proteins(Lin, Schwarz et al. 1997, Molkentin, Lin et al. 1997, Gajewski, Kim et al. 1998, Deepak Srivastava 2000) and the functional impact of these transcriptional factors have further been proved in later embryonic stem cells differentiation into cardiomyocytes studies(Qian and Srivastava 2013, Yi, Mummery et al. 2013).

However, these studies are limited because microarray can only provide information based on EST clones and predicted gene expression, and the relatively low sensitivity and low throughput of microarray in detecting gene expression(Churko, Mantalas et al. 2013, Lappalainen, Sammeth et al. 2013).

Our study, for the first time, explored the real complexity of cardiac transcriptome utilizing the high-throughput technology—deep RNA-Sequencing. During this sequencing effort, we have obtained a total of 335,772,792 reads, among them, 191,049,656 reads can be uniquely mapped to mouse transcriptome. We have identified ~1000 previously un-identified novel exons due to different alternative splicing events. We have also identified >700 lncRNAs that are expressed at significant level in cardiomyocytes. By comparing the pressure-overload induced failing hearts with sham operated normal hearts, we further demonstrated the cardiac transcriptome is dynamically regulated. More than one thousand genes showed differential gene expression and differential alternative splicing in diseased hearts. We also provided evidence, for the first time, that lncRNAs could

have functional impact on cardiac disease based on their expression profile in diseased heart and normal hearts.

A fine-tuned functional genomics in heart

Although the concepts of lncRNA and novel transcripts have been established decades ago, for a long time, these transcriptome components have been poorly studied and considered as transcription by-products or transcription noises. One reason for this is that most lncRNAs and novel splicing variants are not expressed at high level, and the tools to study these transcriptome components are not well established. With the development of high-throughput technology, especially deep RNA-Sequencing, investigators were able to zoom into these previously under-appreciated transcriptome components at single base resolution. Based on the deep sequencing effort in mammal and model systems, it is suggested that, a significant amount of these uncharacterized transcripts are highly conserved across multiple species, and their expression is tightly regulated during different pathological conditions, leading to a new era of " Functional Genomics".(Birney, Stamatoyannopoulos et al. 2007, Mortazavi, Williams et al. 2008, Gerstein, Lu et al. 2010, Muers 2011)

Our study on PKCα-NE further provided strong evidence for the notion of " Functional Genomics". Based on in vivo and in vitro minigene reporter analysis, we have demonstrated this highly conserved novel exon in PKCα is a cardiac specific splicing event tightly regulated by splicing regulator—RBFox1. The novel splicing variant of PKCα has distinct activation and translocation profile comparing to original PKCα upon different hypertrophic stimuli, including PMA,

ISO and AngII. Interestingly, both these two PKCα transcripts induced cardiomyocytes hypertrophic response when overexpressed in vitro, but through different pathways. While the original PKCα interacts with classic chaperon proteins according to our mass spectrometry study, the novel transcript—PKCα-NE interacts specifically with key components of protein synthesis machinery—eEF1A1. In addition, the ratio between these two PKCα transcripts changed during cardiac transcriptome maturation and pathological remodeling.

Taken together, we have used one example-PKCα to demonstrate that, the previously un-annotated splicing variants and lncRNAs can indeed have unique functional impact in cardiac function regulation. The ratio between novel VS original PKCα splicing variant might have different impact mediating cardiomyocytes contractility and hypertrophy response. Establishing a link between PKCα-NE and eEF1A1 can also have therapeutic value in designing inhibitors targeting PKCα activity.

Of course, there are also several limitations in our study that require further investigation.

Firstly, we haven't provided protein evidence for this novel PKCα splicing variant. This novel exon is a very short 48bp exon, and is highly enriched with potential trypsin digestion sites, making the mass spectrometry analysis difficult. Potential solution to this problem is to generate splicing variant specific antibody based on the protein sequence of the novel exon and to identify the existence of this novel exon at protein level during cardiac maturation and disease remodeling.

Secondly, although biochemistry analysis and in vivo gain-of-function study both suggested a significant higher level of PKCα-NE auto-phosphorylation. We have not provided evidence whether the insertion of this novel exon contains novel phosphorylation site, or the insertion of novel exon induces phosphorylation at turn motif. Future analysis would include generating mutant PKCα-NE to replace the potential phosphorylation sites. This would provide detailed molecular mechanism underlying the phosphorylation and activation profile of PKCα-NE.

Lastly, we have demonstrated PKCα-NE interacts and potentially phosphorylates eEF1A1 at least in vitro. But it is not clear whether the phosphorylation event is required for the interaction and whether it is the unique sequence of the novel exon that is required for this interaction. In order to carry out a detailed analysis on this interaction, and to provide insights on pharmacological inhibitor design, mutation analysis based on potential phosphorylation site on eEF1A1 would be necessary and helpful.

A fetal-like alternative splicing profile during pathological remodeling

Earlier section mentioned the importance of key transcription factors regulating cardiac development and maturation. Interestingly, previous studies have also suggested a significant role of these transcription factors during cardiac pathological remodeling. That is, when the mature adult heart undergoes cardiac stress, including hypertrophy and pressure-overload, the silent genes are re-induced, eg. fetal isoform of contractile proteins and fetal type of cardiac ion channels plus the previously mentioned transcription factors including Nkx2.5,

MEF2 ,GATA and TBX proteins (Lynch, Chilibeck et al. 2006, Kuwahara, Nishikimi et al. 2012).

However, studies based on this " fetal gene program" majorly focused on total gene expression regulated at transcription level. Using RNA-Sequencing together with real-time PCR analysis, we have revealed, for the first time, that this fetal reprogramming in heart during cardiac stress is also regulated at alternative splicing level. We have demonstrated a significant number of genes that have shared alternative splicing pattern in neonatal hearts and failing hearts, including genes that are functional important for heart, e.g. MAPK, MEF2, Mfn and CamKinase.

The shared alternative splicing pattern between developing and diseased heart suggested shared molecular regulatory machinery for alternative splicing. Our study has provided evidence that, a cardiac and skeletal muscle enriched splicing regulator—RBFox1 is at least, partially responsible for the fetal like alternative splicing in failing heart. Bioinformatics suggested RBFox1 binding motif is highly enriched among the differentially spliced exons in diseased hearts. We have also demonstrated that this splicing regulator itself is dynamically regulated during cardiac development and pathological remodeling, and directly regulates one such fetal like alternative splicing event in failing heart—MEF2 mutually exclusive alternative splicing.

However, a global fetal like alternative splicing profile in failing heart is more likely to be regulated by multiple players. Our bioinformatics study also suggested there are other binding motifs enriched in the alternatively splicing

events in failing heart. Although RBFox1 is the only splicing regulator that is significantly down-regulated in failing hearts comparing to sham operated hearts according to our RNA-Sequencing analysis, it is important to keep in mind that many splicing regulators are regulated at post-translational level, including Cugbp1(Apponi, Corbett et al. 2011, Dasgupta and Ladd 2012). A detailed loss-of-function study would provide further insights on identification of key splicing regulators responsible for the fetal like alternative splicing during cardiac disease.

A novel regulatory circuit between alternative splicing and transcription

As mentioned earlier, the key component of cardiac development blueprint is transcription factors including MEF2. And the functional significance of these transcription factors is also well established during cardiac hypertrophy and heart failure. Our study, however, focused on transcriptome regulation at alternative splicing level.

The alternative splicing of MEF2 has been documented since last decade as it is a highly conserved mutually exclusive splicing event that is shared by all the MEF2 family members(Bachinski, Sirito et al. 2010). But the functional significance of this alternative splicing event has never been established. Based on our gain-of-function analysis in zebrafish and in vitro cultured cardiomyocytes, the mutually exclusive splicing event of MEF2 generates fetal VS adult splicing variants that have distinct transcription factor activities and target preference. We have further established a link between splicing regulator-RBFox1 and MEF2 alternative splicing. The expression of RBFox1 is dynamically regulated during cardiac transcriptome maturation and pathological remodeling; RBFox1 further mediates

MEF2 mutually exclusive alternative splicing, generating adult or fetal splicing variant; different MEF2 splicing variant have different transcription factor activities toward different set of genes, and further regulate cardiac development and function during different physiological conditions. Our study has established a novel regulatory circuit in cardiac transcriptome, linking alternative splicing and transcriptional regulation, thus provided a fine-tuned regulatory network at different levels in cardiac transcriptome.

For the future study, we would like to provide a more detailed molecular mechanism underlying differential transcription activities between MEF2 alternative splicing variants. Overexpression of MEF2 splicing variants induced different cardiac phenotype in zebrafish, RNA-Sequencing study also suggested different MEF2 splicing variants have different target preferences. The mutually exclusive spliced exon locates right next to MADS box and MEF2 signature, potentially affects MEF2 binding specificity and co-factor interaction(Wu, Dey et al. 2010). Future analysis based on MEF2 different splicing variants ChIP-Seq analysis and immune-precipitation analysis determining the co-factors interaction profile between MEF2 splicing variants would yield a more comprehensive understanding of this splicing event at molecular level.

A novel therapeutic target in heart disease

Our study has identified a cardiac enriched splicing regulator-RBFox1 to be dynamically regulated during cardiac transcriptome maturation and pathological remodeling. Interestingly, the expression level of RBFox1 is tightly regulated under a variety of pathological conditions. We have shown that RBFox1

expression is significantly down-regulated 14 days after severe TAC surgery, but is significantly recovered 1 day post TAC removal. This well correlates with our finding in human samples, where we observed RBFox1 expression is significantly reduced in human dilated cardiomyopathy, but is recovered after LVAD implantation, together with a rescued ejection fraction.

Knocking down zebrafish RBFox1 caused severe cardiac developmental defects and heart failure phenotype; while restored expression of RBFox1 post-TAC is sufficient to preserve mice cardiac function. Taken together, RBFox1 might serve as a potential therapeutic target in heart failure.

Our study further demonstrated a critical role of MEF2 alternative splicing in RBFox1 mediated cardiac splicing regulation based on our zebrafish double-knockdown studies and cardiomyocytes studies. Inhibitor targeting to either RBFox1 or MEF2 different splicing variants would provide therapeutic value in both hypertrophy and heart failure.

Future Prospect

Our study had made several interesting discoveries providing a more comprehensive understanding of the total cardiac transcriptome complexity. Firstly, we have identified additional component of cardiac transcriptome. Future study in this area will involve carrying out large scale proteomics study to identify the existence and protein expression level of these novel transcripts and splicing variants in normal and diseased hearts. In order to identify the functional impact of these un-explored novel transcriptome components in heart, a high-throughput

RNAi screening will also reveal their roles mediating cardiomyocytes contractility and function.

Our finding based on the novel cardiac splicing event in PKCα also pointed to potential novel cardiac splicing regulators. Our study has demonstrated in vivo and in vitro that RBFox1 is at least partially regulating PKCα alternative splicing. However, there are additional highly conserved cis-regulatory motifs adjacent to PKCα-novel exon. Unbiased study using RNA-immune-precipitation would provide additional insights into the cardiac specific splicing regulator candidates.

Lastly, our deep RNA-Sequencing analysis, together with high-throughput studies that reported by other groups, have provided strong evidence that there are many components of cardiac transcriptome that contributing to the total transcriptome complexity and dynamics during disease. Our study used PKCα-NE as one example to demonstrate that these previously uncharacterized components could indeed functionally important. Future study focusing on functional characterization based on the novel transcript clusters and lncRNA would also generate additional insights towards a better understanding of the functional genomics in heart.

References:

Adamidi, C. (2011). "De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics." Genome Res. **21**: 1193-1200.

Apponi, L. H., A. H. Corbett and G. K. Pavlath (2011). "RNA-binding proteins and gene regulation in myogenesis." Trends in Pharmacological Sciences **32**(11): 652-658.

Au, K. F., H. Jiang, L. Lin, Y. Xing and W. H. Wong (2010). "Detection of splice junctions from paired-end RNA-seq data by SpliceMap." Nucleic Acids Res. **38**: 4570-4578.

Bachinski, L. L., M. Sirito, M. Böhme, K. A. Baggerly, B. Udd and R. Krahe (2010). "Altered MEF2 isoforms in myotonic dystrophy and other neuromuscular disorders." Muscle & Nerve **42**(6): 856-863.

Barry, S. P., S. M. Davidson and P. A. Townsend (2008). "Molecular regulation of cardiac hypertrophy." Int J Biochem Cell Biol **40**(10): 2023-2039.

Barry, S. P., S. M. Davidson and P. A. Townsend (2008). "Molecular regulation of cardiac hypertrophy." The International Journal of Biochemistry & Cell Biology **40**(10): 2023-2039.

Bayer AL, H. M., Patel N, Porter M, Engman S, Samarel AM (2003). "Alterations in protein kinase C isoenzyme expression and autophosphorylation during the

progression of pressure-overload induced left ventricular hypertrophy." Mol Cell Biochem **243**(2): 145-152.

Birney, E., J. Stamatoyannopoulos, A. Dutta, R. Guigo, T. Gingeras, E. Margulies, Z. Weng, M. Snyder, E. Dermitzakis, J. Stamatoyannopoulos, R. Thurman, M. Kuehn, C. Taylor, S. Neph, C. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. Greenbaum, R. Andrews, P. Flicek, P. Boyle, H. Cao, N. Carter, G. Clelland, S. Davis, N. Day, P. Dhami, S. Dillon and M. Dorschner (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." Nature **447**: 799 - 816.

Birol, I. (2009). "De novo transcriptome assembly with ABySS." Bioinformatics **25**: 2872-2877.

Bland, C. S., E. T. Wang, A. Vu, M. P. David, J. C. Castle, J. M. Johnson, C. B. Burge and T. A. Cooper (2010). "Global regulation of alternative splicing during myogenic differentiation." Nucleic Acids Res **38**(21): 7651-7664.

Boutz, P. L., P. Stoilov, Q. Li, C.-H. Lin, G. Chawla, K. Ostrow, L. Shiue, M. Ares and D. L. Black (2007). "A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons." Genes & Development **21**(13): 1636-1652.

Buljan, M., G. Chalancon, S. Eustermann, G. P. Wagner, M. Fuxreiter, A. Bateman and M. M. Babu (2012). "Tissue-Specific Splicing of Disordered Segments that Embed Binding Motifs Rewires Protein Interaction Networks." Molecular Cell **46**(6): 871-883.

Chakraborti, S., S. Roy, A. Chowdhury, A. Mandal and T. Chakraborti (2013). "Role of PKCα−p38MAPK−Giα axis in peroxynitrite-mediated inhibition of β-adrenergic response in pulmonary artery smooth muscle cells." Cellular Signalling **25**(2): 512-526.

Chang, S., T. A. McKinsey, C. L. Zhang, J. A. Richardson, J. A. Hill and E. N. Olson (2004). "Histone Deacetylases 5 and 9 Govern Responsiveness of the Heart to a Subset of Stress Signals and Play Redundant Roles in Heart Development." Molecular and Cellular Biology **24**(19): 8467-8476.

Chen Gao, Y. W. (2012). "Global impact of RNA splicing on transcriptome remodeling in the heart." Journal of Zhejiang University Science B **13**(8): 603-608.

Chen, M. and J. L. Manley (2009). "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." Nat Rev Mol Cell Biol **10**(11): 741-754.

Churko, J. M., G. L. Mantalas, M. P. Snyder and J. C. Wu (2013). "Overview of High Throughput Sequencing Technologies to Elucidate Molecular Pathways in Cardiovascular Diseases." Circulation Research **112**(12): 1613-1623.

Dasgupta, T. and A. N. Ladd (2012). "The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins." Wiley Interdisciplinary Reviews: RNA **3**(1): 104-121.

de la Grange, P., L. Gratadou, M. Delord, M. Dutertre and D. Auboeuf (2010). "Splicing factor and exon profiling across human tissues." Nucleic Acids Research **38**(9): 2825-2838.

De Windt, L. J., H. W. Lim, S. Haq, T. Force and J. D. Molkentin (2000). "Calcineurin Promotes Protein Kinase C and c-Jun NH2-terminal Kinase Activation in the Heart: CROSS-TALK BETWEEN CARDIAC HYPERTROPHIC SIGNALING PATHWAYS." Journal of Biological Chemistry **275**(18): 13571-13579.

Deepak Srivastava, E. N. O. (2000). "A genetic blueprint for cardiac development." Nature **407**: 221-226.

Ding, J.-H., X. Xu, D. Yang, P.-H. Chu, N. D. Dalton, Z. Ye, J. M. Yeakley, H. Cheng, R.-P. Xiao, J. Ross, J. Chen and X.-D. Fu (2004). "Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart." EMBO J **23**(4): 885-896.

Durocher, D., F. Charron, R. Warren, R. J. Schwartz and M. Nemer (1997). "The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors." EMBO J. **16**: 5687-5696.

Edwards, A. S. and A. C. Newton (1997). "Phosphorylation at Conserved Carboxyl-terminal Hydrophobic Motif Regulates the Catalytic and Regulatory Domains of Protein Kinase C." Journal of Biological Chemistry **272**(29): 18382-18390.

Eric Chruchill, G. B., Alice Vallentin, Tomoyoshi Koyanagi, and Daria Mochly-Rosen (2008). "PKC Isozymes in Chronic Cardiac Disease: Possible Therapeutic Targets." Annual Reviews Pharmacology and Toxicology(48): 569-599.

Gajewski, K., Y. Kim, Y. M. Lee, E. N. Olson and R. A. Schulz (1998). "D-Mef2: a target for tinman activation during Drosophila heart development." EMBO J. **16**: 515-522.

Gang, H., Y. Hai, R. Dhingra, J. W. Gordon, N. Yurkova, Y. Aviv, H. Li, F. Aguilar, A. Marshall, E. Leygue and L. A. Kirshenbaum (2011). "A Novel Hypoxia-Inducible Spliced Variant of Mitochondrial Death Gene Bnip3 Promotes Survival of Ventricular Myocytes / Novelty and Significance." Circulation Research **108**(9): 1084-1092.

Gao, G. and S. C. Dudley Jr "RBM25/LUC7L3 Function in Cardiac Sodium Channel Splicing Regulation of Human Heart Failure." Trends in Cardiovascular Medicine(0).

Gao, G. and S. C. Dudley Jr (2013). "RBM25/LUC7L3 Function in Cardiac Sodium Channel Splicing Regulation of Human Heart Failure." Trends in Cardiovascular Medicine **23**(1): 5-8.

Gerstein, M. B., Z. J. Lu, E. L. Van Nostrand, C. Cheng, B. I. Arshinoff, T. Liu, K. Y. Yip, R. Robilotto, A. Rechtsteiner, K. Ikegami, P. Alves, A. Chateigner, M. Perry, M. Morris, R. K. Auerbach, X. Feng, J. Leng, A. Vielle, W. Niu, K. Rhrissorrakrai, A. Agarwal, R. P. Alexander, G. Barber, C. M. Brdlik, J. Brennan, J. J. Brouillet, A. Carr, M.-S. Cheung, H. Clawson, S. Contrino, L. O. Dannenberg, A. F. Dernburg, A. Desai, L. Dick, A. C. Dosé, J. Du, T. Egelhofer, S. Ercan, G. Euskirchen, B. Ewing, E. A. Feingold, R. Gassmann, P. J. Good, P. Green, F. Gullier, M. Gutwein, M. S. Guyer, L. Habegger, T. Han, J. G. Henikoff, S. R. Henz, A. Hinrichs, H. Holster, T. Hyman, A. L. Iniguez, J. Janette, M.

Jensen, M. Kato, W. J. Kent, E. Kephart, V. Khivansara, E. Khurana, J. K. Kim, P. Kolasinska-Zwierz, E. C. Lai, I. Latorre, A. Leahey, S. Lewis, P. Lloyd, L. Lochovsky, R. F. Lowdon, Y. Lubling, R. Lyne, M. MacCoss, S. D. Mackowiak, M. Mangone, S. McKay, D. Mecenas, G. Merrihew, D. M. Miller, A. Muroyama, J. I. Murray, S.-L. Ooi, H. Pham, T. Phippen, E. A. Preston, N. Rajewsky, G. Rätsch, H. Rosenbaum, J. Rozowsky, K. Rutherford, P. Ruzanov, M. Sarov, R. Sasidharan, A. Sboner, P. Scheid, E. Segal, H. Shin, C. Shou, F. J. Slack, C. Slightam, R. Smith, W. C. Spencer, E. O. Stinson, S. Taing, T. Takasaki, D. Vafeados, K. Voronina, G. Wang, N. L. Washington, C. M. Whittle, B. Wu, K.-K. Yan, G. Zeller, Z. Zha, M. Zhong, X. Zhou, m. Consortium, J. Ahringer, S. Strome, K. C. Gunsalus, G. Micklem, X. S. Liu, V. Reinke, S. K. Kim, L. W. Hillier, S. Henikoff, F. Piano, M. Snyder, L. Stein, J. D. Lieb and R. H. Waterston (2010). "Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project." Science **330**(6012): 1775-1787.

Greulich, F., C. Rudat and A. Kispert (2011). "Mechanisms of T-box gene function in the developing heart." Cardiovascular Research **91**(2): 212-222.

Guo, W., S. Schafer, M. L. Greaser, M. H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A. M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. A. Hacker, K. W. Saupe, G. W. Dec, P. T. Ellinor, C. A. MacRae, B. Spallek, R. Fischer, A. Perrot, C. Ozcelik, K. Saar, N. Hubner and M. Gotthardt (2012). "RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing." Nat Med **18**(5): 766-773.

Gusterson, R. J., E. Jazrawi, I. M. Adcock and D. S. Latchman (2003). "The Transcriptional Co-activators CREB-binding Protein (CBP) and p300 Play a Critical Role in Cardiac Hypertrophy That Is Dependent on Their Histone Acetyltransferase Activity." Journal of Biological Chemistry **278**(9): 6838-6847.

Haarberg, K. M. K., J. Li, J. Heinrichs, D. Wang, C. Liu, C. C. Bronk, K. Kaosaard, A. M. Owyang, S. Holland, E. Masuda, K. Tso, B. R. Blazar, C. Anasetti, A. A. Beg and X.-Z. Yu (2013). "Pharmacologic inhibition of PKCα and PKCθ prevents GVHD while preserving GVL activity in mice." Blood **122**(14): 2500-2511.

Hahn, H. S., Y. Marreez, A. Odley, A. Sterbling, M. G. Yussman, K. C. Hilty, I. Bodi, S. B. Liggett, A. Schwartz and G. W. Dorn (2003). "Protein Kinase Cα Negatively Regulates Systolic and Diastolic Function in Pathological Hypertrophy." Circulation Research **93**(11): 1111-1119.

Hallegger, M., M. Llorian and C. W. J. Smith (2010). "Alternative splicing: global insights." FEBS Journal **277**(4): 856-866.

Hambleton, M., H. Hahn, S. T. Pleger, M. C. Kuhn, R. Klevitsky, A. N. Carr, T. F. Kimball, T. E. Hewett, G. W. Dorn, W. J. Koch and J. D. Molkentin (2006). "Pharmacological- and Gene Therapy-Based Inhibition of Protein Kinase Cα/β Enhances Cardiac Contractility and Attenuates Heart Failure." Circulation **114**(6): 574-582.

Hamplova B, N. F., Kolar F, Novakova O (2010). "Transient upregulation of protein kinase C in pressure-overloaded neonatal rat myocardium
." Physiological Research **59**(1): 25-33.

Harvey, R. P. (1996). Dev. Biol. **178**: 203-216.

Henderson, D. J. and A. J. Copp (1998). "Versican Expression Is Associated With Chamber Specification, Septation, and Valvulogenesis in the Developing Mouse Heart." Circulation Research **83**(5): 523-532.

Herrera, M., G. B. Silva and J. L. Garvin (2010). "Angiotensin II Stimulates Thick Ascending Limb Superoxide Production via Protein Kinase Cα-dependent NADPH Oxidase Activation." Journal of Biological Chemistry **285**(28): 21323-21328.

Honda, A., Y. Valogne, M. Bou Nader, C. Br√©chot and J. Faivre (2012). "An Intron-Retaining Splice Variant of Human Cyclin A2, Expressed in Adult Differentiated Tissues, Induces a G1/S Cell Cycle Arrest In Vitro." PLoS ONE **7**(6): e39249.

Hong Ki Song, S.-E. H., Taeyong Kim, Do Han Kim (2012). "Deep RNA Sequencing Reveals Novel Cardiac Transcriptomic Signatures for Physiological and Pathological Hypertrophy." PLOS ONE.

Huang, Z.-P., J. Chen, H. Y. Seok, Z. Zhang, M. Kataoka, X. Hu and D.-Z. Wang (2013). "MicroRNA-22 Regulates Cardiac Hypertrophy and Remodeling in Response to Stress." Circulation Research **112**(9): 1234-1243.

Itoh, S., B. Ding, C. P. Bains, N. Wang, Y. Takeishi, T. Jalili, G. L. King, R. A. Walsh, C. Yan and J.-i. Abe (2005). "Role of p90 Ribosomal S6 Kinase (p90RSK) in Reactive Oxygen Species and Protein Kinase C β (PKC-β)-mediated Cardiac Troponin I Phosphorylation." Journal of Biological Chemistry **280**(25): 24135-24142.

Jin, Y., H. Suzuki, S. Maegawa, H. Endo, S. Sugano, K. Hashimoto, K. Yasuda and K. Inoue (2003). "A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG." EMBO J **22**(4): 905-912.

Jingjing, L., Y. Xue, N. Agarwal and R. S. Roque (1999). "Human Müller cells express VEGF183, a novel spliced variant of vascular endothelial growth factor." Investigative Ophthalmology & Visual Science **40**(3): 752-759.

Joanna Kobylinska, W. D., Monika Cendrowska-Pinkosz, Anna Dworzanska, Teresa Hermanowics-Dryka, Joanna Kiszka, Elzbieta Staroslawska, Franciszek Burdan (2013). "Morphological and Moelcular Bases of cardiac development." Postepy Hig Med Dosw.

Jr, G. G. a. S. C. D. (2013). "SCN5A splicing variants and the possibility of predicting heart failure-associated arrhythmia." Expert Reveiw of Cardiovascular Therapy **11**(2): 117-119.

Kalsotra, A., K. Wang, P.-F. Li and T. A. Cooper (2010). "MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development." Genes & Development **24**(7): 653-658.

Kilic, A., A. Velic, L. J. De Windt, L. Fabritz, M. Voss, D. Mitko, M. Zwiener, H. A. Baba, M. van Eickels, E. Schlatter and M. Kuhn (2005). "Enhanced Activity of the Myocardial Na+/H+ Exchanger NHE-1 Contributes to Cardiac Remodeling in Atrial Natriuretic Peptide Receptor–Deficient Mice." Circulation **112**(15): 2307-2317.

Kirchhefer, U., A. Heinick, S. König, T. Kristensen, F. U. Müller, M. D. Seidl and P. Boknik (2013). "Protein phosphatase 2A is regulated by PKCα-dependent

phosphorylation of its targeting subunit B56α at Ser41." Journal of Biological Chemistry.

Koide, Y., K. Tamura, A. Suzuki, K. Kitamura, K. Yokoyama, T. Hashimoto, N. Hirawa, M. Kihara, Ohno, Shigeo and S. Umemura (2003). "<B>Differential Induction of Protein Kinase C Isoforms at the Cardiac Hypertrophy Stage and Congestive Heart Failure Stage in Dahl Salt-Sensitive Rats</B>." Hypertension Research **26**(5): 421-426.

Koshelev, M., S. Sarma, R. E. Price, X. H. T. Wehrens and T. A. Cooper (2010). "Heart-specific overexpression of CUGBP1 reproduces functional and molecular abnormalities of myotonic dystrophy type 1." Human Molecular Genetics **19**(6): 1066-1075.

Kuwahara, K., T. Nishikimi and K. Nakao (2012). "Transcriptional Regulation of the Fetal Cardiac Gene Program." Journal of Pharmacological Sciences **119**(3): 198-203.

Kuwahara, K., T. Nishikimi and K. Nakao (2012). "Transcriptional regulation of the fetal cardiac gene program." J Pharmacol Sci **119**(3): 198-203.

Ladage, D., L. Tilemann, K. Ishikawa, R. N. Correll, Y. Kawase, S. R. Houser, J. D. Molkentin and R. J. Hajjar (2011). "Inhibition of PKCα/β With Ruboxistaurin Antagonizes Heart Failure in Pigs After Myocardial Infarction Injury." Circulation Research **109**(12): 1396-1400.

Lappalainen, T., M. Sammeth, M. R. Friedlander, P. A. C. t Hoen, J. Monlong, M. A. Rivas, M. Gonzalez-Porta, N. Kurbatova, T. Griebel, P. G. Ferreira, M. Barann, T. Wieland, L. Greger, M. van Iterson, J. Almlof, P. Ribeca, I.

Pulyakhina, D. Esser, T. Giger, A. Tikhonov, M. Sultan, G. Bertier, D. G. MacArthur, M. Lek, E. Lizano, H. P. J. Buermans, I. Padioleau, T. Schwarzmayr, O. Karlberg, H. Ongen, H. Kilpinen, S. Beltran, M. Gut, K. Kahlem, V. Amstislavskiy, O. Stegle, M. Pirinen, S. B. Montgomery, P. Donnelly, M. I. McCarthy, P. Flicek, T. M. Strom, C. The Geuvadis, H. Lehrach, S. Schreiber, R. Sudbrak, A. Carracedo, S. E. Antonarakis, R. Hasler, A.-C. Syvanen, G.-J. van Ommen, A. Brazma, T. Meitinger, P. Rosenstiel, R. Guigo, I. G. Gut, X. Estivill and E. T. Dermitzakis (2013). "Transcriptome and genome sequencing uncovers functional variation in humans." Nature **501**(7468): 506-511.

Lara-Pezzi, E., J. Gómez-Salinero, A. Gatto and P. García-Pavía (2013). "The Alternative Heart: Impact of Alternative Splicing in Heart Disease." Journal of Cardiovascular Translational Research **6**(6): 945-955.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research.

Lee, J.-H., C. Gao, G. Peng, C. Greer, S. Ren, Y. Wang and X. Xiao (2011). "Analysis of Transcriptome Complexity Through RNA Sequencing in Normal and Failing Murine Hearts." Circulation Research **109**(12): 1332-1341.

Li, S., W. Guo, B. M. Schmitt and M. L. Greaser (2012). "Comprehensive analysis of titin protein isoform and alternative splicing in normal and mutant rats." Journal of Cellular Biochemistry **113**(4): 1265-1273.

Lin, Q., J. Schwarz, C. Bucana and E. N. Olson (1997). "Control of Mouse Cardiac Morphogenesis and Myogenesis by Transcription Factor MEF2C." Science **276**(5317): 1404-1407.

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

Linke, W. A. and S. Bucker (2012). "King of hearts: a splicing factor rules cardiac proteins." Nat Med **18**(5): 660-661.

Liu, Q., X. Chen, S. M. MacDonnell, E. G. Kranias, J. N. Lorenz, M. Leitges, S. R. Houser and J. D. Molkentin (2009). "Protein Kinase Cα, but Not PKCβ or PKCγ, Regulates Contractility and Heart Failure Susceptibility: Implications for Ruboxistaurin as a Novel Therapeutic Approach." Circulation Research **105**(2): 194-200.

Lu, G., S. Ren, P. Korge, J. Choi, Y. Dong, J. Weiss, C. Koehler, J.-n. Chen and Y. Wang (2007). "A novel mitochondrial matrix serine/threonine protein phosphatase regulates the mitochondria permeability transition pore and is essential for cellular survival and development." Genes & Development **21**(7): 784-796.

Lu, G., H. Sun, P. She, J.-Y. Youn, S. Warburton, P. Ping, T. M. Vondriska, H. Cai, C. J. Lynch and Y. Wang (2009). "Protein phosphatase 2Cm is a critical regulator of branched-chain amino acid catabolism in mice and cultured cells." The Journal of Clinical Investigation **119**(6): 1678-1687.

Lu, J., T. A. McKinsey, R. L. Nicol and E. N. Olson (2000). "Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases." Proceedings of the National Academy of Sciences **97**(8): 4070-4075.

Lu, J., T. A. McKinsey, C.-L. Zhang and E. N. Olson (2000). "Regulation of Skeletal Myogenesis by Association of the MEF2 Transcription Factor with Class II Histone Deacetylases." Molecular Cell **6**(2): 233-244.

Lynch, J. M., K. Chilibeck, Y. Qui and M. Michalak (2006). "Assembling Pieces of the Cardiac Puzzle; Calreticulin and Calcium-Dependent Pathways in Cardiac Development, Health, and Disease." Trends in Cardiovascular Medicine **16**(3): 65-69.

Manetti, M., S. Guiducci, L. Ibba-Manneschi and M. Matucci-Cerinic (2011). "Impaired Angiogenesis in Systemic Sclerosis: The Emerging Role of the Antiangiogenic VEGF165b Splice Variant." Trends in Cardiovascular Medicine **21**(7): 204-210.

Martin, J. A. and Z. Wang (2011). "Next-generation transcriptome assembly." Nat Rev Genet **12**(10): 671-682.

McManus, C. J. and B. R. Graveley (2011). "RNA structure and the mechanisms of alternative splicing." Current Opinion in Genetics & Development **21**(4): 373-379.

Merdzhanova, G., S. Gout, M. Keramidas, V. Edmond, J. L. Coll, C. Brambilla, E. Brambilla, S. Gazzeri and B. Eymin (2010). "The transcription factor E2F1 and the SR protein SC35 control the ratio of pro-angiogenic versus antiangiogenic

isoforms of vascular endothelial growth factor-A to inhibit neovascularization in vivo." Oncogene **29**(39): 5392-5403.

Mjaatvedt, C. H., H. Yamamura, A. A. Capehart, D. Turner and R. R. Markwald (1998). "TheCspg2Gene, Disrupted in thehdfMutant, Is Required for Right Cardiac Chamber and Endocardial Cushion Formation." Developmental Biology **202**(1): 56-66.

Molkentin, J., Q. Lin, S. A. Duncan and E. N. Olson (1997). "Requirement of the GATA4 transcription factor for heart tube formation and ventral morphogenesis." Genes Dev. **11**: 1061-1072.

Molnár, A., A. Borbély, D. Czuriga, S. M. Ivetta, S. Szilágyi, Z. Hertelendi, E. T. Pásztor, Á. Balogh, Z. Galajda, T. Szerafin, K. Jaquet, Z. Papp, I. Édes and A. Tóth (2009). "Protein Kinase C Contributes to the Maintenance of Contractile Force in Human Ventricular Cardiomyocytes." Journal of Biological Chemistry **284**(2): 1031-1039.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer and B. Wold (2008). "Mapping and quantifying mammalian transcriptomes by RNA-seq." Nature Methods **5**: 621-628.

Muers, M. (2011). "Functional genomics: The modENCODE guide to the genome." Nat Rev Genet **12**(2): 80-80.

Murphy, L. L., A. J. Moon-Grady, B. F. Cuneo, R. T. Wakai, S. Yu, J. D. Kunic, D. W. Benson and A. L. George Jr (2012). "Developmentally regulated SCN5A splice variant potentiates dysfunction of a novel mutation associated with severe fetal arrhythmia." Heart Rhythm **9**(4): 590-597.

Nalini Raghavachari, J. B., Yanqin Yang, Poching Liu, Kimberly Woodhouse, Daniel Levy, Chris O'Donnel, Peter J Munson and Gregory Kato (2012). "A systematic comparison and evaluation of high density exon arrays and RNA-seq technology used to unravel the peripheral blood transcriptome of sickle cell disease." BMC Medical Genomics **5**(28).

Olson, E. (2002). "A genetic blueprint for growth and development of the heart." Harvey Lect **98**: 41-64.

Olson, E. N. (2006). "Gene regulatory networks in the evolution and development of the heart." Science **313**(5795): 1922-1927.

Orenes-Piñero, E., S. Montoro-García, J. V. Patel, M. Valdés, F. Marín and G. Y. H. Lip (2013). "Role of microRNAs in cardiac remodelling: New insights and future perspectives." International Journal of Cardiology **167**(5): 1651-1659.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey and B. J. Blencowe (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nat Genet **40**(12): 1413-1415.

Pass, J. M., J. Gao, W. K. Jones, W. B. Wead, X. Wu, J. Zhang, C. P. Baines, R. Bolli, Y.-T. Zheng, I. G. Joshua and P. Ping (2001). "Enhanced PKCβII translocation and PKCβII-RACK1 interactions in PKCε-induced heart failure: a role for RACK1." American Journal of Physiology - Heart and Circulatory Physiology **281**(6): H2500-H2510.

Ping, P., J. Zhang, Y. Qiu, X.-L. Tang, S. Manchikalapudi, X. Cao and R. Bolli (1997). "Ischemic Preconditioning Induces Selective Translocation of Protein Kinase C Isoforms ε and η in the Heart of Conscious Rabbits Without Subcellular

Redistribution of Total Protein Kinase C Activity." Circulation Research **81**(3): 404-414.

Qian, L., Y. Huang, C. I. Spencer, A. Foley, V. Vedantham, L. Liu, S. J. Conway, J.-d. Fu and D. Srivastava (2012). "In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes." Nature **advance online publication**.

Qian, L. and D. Srivastava (2013). "Direct Cardiac Reprogramming: From Developmental Biology to Cardiac Regeneration." Circulation Research **113**(7): 915-921.

Ramchatesingh, J., A. Zahler, K. Neugebauer, M. Roth and T. Cooper (1995). "A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer." Mol. Cell. Biol. **15**(9): 4898-4907.

Rana, M. S., V. M. Christoffels and A. F. M. Moorman (2013). "A molecular and genetic outline of cardiac morphogenesis." Acta Physiologica **207**(4): 588-615.

Refaat, M. M., S. A. Lubitz, S. Makino, Z. Islam, J. M. Frangiskakis, H. Mehdi, R. Gutmann, M. L. Zhang, H. L. Bloom, C. A. MacRae, S. C. Dudley, A. A. Shalaby, R. Weiss, D. M. McNamara, B. London and P. T. Ellinor (2012). "Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy." Heart rhythm : the official journal of the Heart Rhythm Society **9**(3): 390-396.

Rybin, V. O., X. Xu and S. F. Steinberg (1999). "Activated Protein Kinase C Isoforms Target to Cardiomyocyte Caveolae: Stimulation of Local Protein Phosphorylation." Circulation Research **84**(9): 980-988.

Severino, P., L. Oliveira, N. Torres, F. Andreghetto, M. de Fatima Guarizo Klingbeil, R. Moyses, V. Wunsch-Filho, F. Nunes, M. Mathor, A. Paschoal and A. Durham (2013). "High-throughput sequencing of small RNA transcriptomes reveals critical biological features targeted by microRNAs in cell models used for squamous cell cancer research." BMC Genomics **14**(1): 735.

Sharma GD, N. H., Antonov AS, Gerrity RG, von Geldern T, Pandey KN (2002). "Expression of atrial natriuretic peptide receptor-A antagonizes the mitogen-activated protein kinases (Erk2 and P38MAPK) in cultured human vascular smooth muscle cells." Mol Cell Biochem **233**(2-3): 165-173.

Song, K., Y.-J. Nam, X. Luo, X. Qi, W. Tan, G. N. Huang, A. Acharya, C. L. Smith, M. D. Tallquist, E. G. Neilson, J. A. Hill, R. Bassel-Duby and E. N. Olson (2012). "Heart repair by reprogramming non-myocytes with cardiac transcription factors." Nature **485**(7400): 599-604.

Srivastava, D. (1997). "Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND." Nature Genet. **16**: 154-160.

Srivastava, D., P. Cserjesi and E. N. Olson (1995). "New subclass of bHLH proteins required for cardiac morphogenesis." Science **270**: 1995-1999.

Steinberg, S. F. (2008). "Structural Basis of Protein Kinase C Isoform Function." Physiological Reviews **88**(4): 1341-1378.

Sultan, M., M. H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf, M. Seifert, T. Borodina, A. Soldatov, D. Parkhomchuk, D. Schmidt, S. O'Keeffe, S. Haas, M. Vingron, H. Lehrach and M.-L. Yaspo (2008). "A Global View of Gene

Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome." Science **321**(5891): 956-960.

Sun, W., X. You, A. Gogol-Doring, H. He, Y. Kise, M. Sohn, T. Chen, A. Klebes, D. Schmucker and W. Chen (2013). "Ultra-deep profiling of alternatively spliced Drosophila Dscam isoforms by circularization-assisted multi-segment sequencing." EMBO J **32**(14): 2029-2038.

Taegtmeyer, H., S. Sen and D. Vela (2010). "Return to the fetal gene program." Annals of the New York Academy of Sciences **1188**(1): 191-198.

Tischer, E., D. Gospodarowicz, R. Mitchell, M. Silva, J. Schilling, K. Lau, T. Crisp, J. C. Fiddes and J. A. Abraham (1989). "Vascular endothelial growth factor: A new member of the platelet-derived growth factor gene family." Biochemical and Biophysical Research Communications **165**(3): 1198-1206.

Turbendian, H. K., M. Gordillo, S.-Y. Tsai, J. Lu, G. Kang, T.-C. Liu, A. Tang, S. Liu, G. I. Fishman and T. Evans (2013). "GATA factors efficiently direct cardiac fate from embryonic stem cells." Development **140**(8): 1639-1644.

Ucar, A., S. K. Gupta, J. Fiedler, E. Erikci, M. Kardasinski, S. Batkai, S. Dangwal, R. Kumarswamy, C. Bang, A. Holzmann, J. Remke, M. Caprio, C. Jentzsch, S. Engelhardt, S. Geisendorf, C. Glas, T. G. Hofmann, M. Nessling, K. Richter, M. Schiffer, L. Carrier, L. C. Napp, J. Bauersachs, K. Chowdhury and T. Thum (2012). "The miRNA-212/132 family regulates both cardiac hypertrophy and cardiomyocyte autophagy." Nat Commun **3**: 1078.

Wahbi, K., V. Algalarrondo, H. M. Bécane, V. Fressart, C. Beldjord, K. Azibi, A. Lazarus, N. Berber, H. Radvanyi-Hoffman, T. Stojkovic, A. Béhin, P. Laforêt, B.

Eymard, S. Hatem and D. Duboc "Brugada syndrome and abnormal splicing of SCN5A in myotonic dystrophy type 1." Archives of Cardiovascular Diseases(0).

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth and C. B. Burge (2008). "Alternative isoform regulation in human tissue transcriptomes." Nature **456**(7221): 470-476.

Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nat Rev Genet **10**(1): 57-63.

Warf, M. B. and J. A. Berglund (2007). "MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T." RNA **13**(12): 2238-2251.

Wei Guo, S. J. B., Karla Esbona and Marion L. Greaser (2010). "Titin Diversity-Alternative Splicing Gone Wild." J Biomed Biotechnol.

Weikard, R., F. Hadlich and C. Kuehn (2013). "Identification of novel transcripts and noncoding RNAs in bovine skin by deep next generation sequencing." BMC Genomics **14**(1): 789.

Westerfield, M. (1995). "The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio)." Eugene: University of Oregon.

Witten, J. T. and J. Ule (2011). "Understanding splicing regulation through RNA splicing maps." Trends in Genetics **27**(3): 89-97.

Wu, Y., R. Dey, A. Han, N. Jayathilaka, M. Philips, J. Ye and L. Chen (2010). "Structure of the MADS-box/MEF2 Domain of MEF2A Bound to DNA and Its Implication for Myocardin Recruitment." Journal of Molecular Biology **397**(2): 520-533.

Xiao, X., Z. Wang, M. Jang and C. B. Burge (2007). "Coevolutionary networks of splicing cis-regulatory elements." Proceedings of the National Academy of Sciences **104**(47): 18583-18588.

Xu, X., D. Yang, J.-H. Ding, W. Wang, P.-H. Chu, N. D. Dalton, H.-Y. Wang, J. R. Bermingham Jr, Z. Ye, F. Liu, M. G. Rosenfeld, J. L. Manley, J. Ross Jr, J. Chen, R.-P. Xiao, H. Cheng and X.-D. Fu (2005). "ASF/SF2-Regulated CaMKII[delta] Alternative Splicing Temporally Reprograms Excitation-Contraction Coupling in Cardiac Muscle." Cell **120**(1): 59-72.

Yae, T., K. Tsuchihashi, T. Ishimoto, T. Motohara, M. Yoshikawa, G. J. Yoshida, T. Wada, T. Masuko, K. Mogushi, H. Tanaka, T. Osawa, Y. Kanki, T. Minami, H. Aburatani, M. Ohmura, A. Kubo, M. Suematsu, K. Takahashi, H. Saya and O. Nagano (2012). "Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell." Nat Commun **3**: 883.

Yanazume, T., K. Hasegawa, T. Morimoto, T. Kawamura, H. Wada, A. Matsumori, Y. Kawase, M. Hirai and T. Kita (2003). "Cardiac p300 Is Involved in Myocyte Growth with Decompensated Heart Failure." Molecular and Cellular Biology **23**(10): 3593-3606.

Yi, B. A., C. L. Mummery and K. R. Chien (2013). "Direct Cardiomyocyte Reprogramming: A New Direction for Cardiovascular Regenerative Medicine." Cold Spring Harbor Perspectives in Medicine **3**(9).

Zhang, C. L., T. A. McKinsey, S. Chang, C. L. Antos, J. A. Hill and E. N. Olson (2002). "Class II Histone Deacetylases Act as Signal-Responsive Repressors of Cardiac Hypertrophy." Cell **110**(4): 479-488.

Zhang, S.-S. and R. M. Shaw (2013). "Multilayered regulation of cardiac ion channels." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1833**(4): 876-885.

Zheng, S., R. Damoiseaux, L. Chen and D. L. Black (2013). "A broadly applicable high-throughput screening strategy identifies new regulators of Dlg4 (Psd-95) alternative splicing." Genome Research **23**(6): 998-1007.