

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Utilizing Systems Genetics Approaches to Identify Novel Molecular Mechanisms in Cardiovascular Diseases

**Permalink**

<https://escholarship.org/uc/item/7hm799d2>

**Author**

Romay, Milagros De La Caridad

**Publication Date**

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Utilizing Systems Genetics Approaches to Identify Novel Molecular Mechanisms in  
Cardiovascular Diseases

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of  
Philosophy in Microbiology, Immunology and Molecular Genetics

by

Milagros De La Caridad Romay

2016

© Copyright by

Milagros De La Caridad Romay

2016

## ABSTRACT OF THE DISSERTATION

Utilizing Systems Genetics Approaches to Identify Novel Molecular Mechanisms in  
Cardiovascular Diseases

by

Milagros De La Caridad Romay

Doctor of Philosophy in Microbiology, Immunology and Molecular Genetics

University of California, Los Angeles, 2016

Professor Aldons J. Lusis, Chair

Despite the success of focused, reductionist approaches in characterizing the pathophysiology of cardiovascular diseases (CVDs), current estimates predict that 24 million deaths annually will be due to CVDs by 2030. Emphasizing the use of genetic variation in combination with mathematical modeling and integration of next generation –omics profiling technologies, systems genetics characterizes the flow of biological information in physiologic and pathologic states to allow investigators to understand the molecular interactions in a system. To assess the feasibility of systems genetics based methodologies in identifying novel interactions that contribute to CVD pathology I have used a combination of animal and cell culture models to recapitulate key processes involved in two CVDs: atherosclerosis and congestive heart failure.

Atherosclerosis, a systemic disorder characterized by the narrowing of arteries is the underlying cause of the majority of clinical cardiovascular events. Formation of the atherosclerotic plaque is driven by chronic endothelial activation from exposure to oxidized phospholipids that accumulate within the vessel wall. Using systems approaches we have identified miRs-21-3p and -27a-5p as novel regulators of NF- $\kappa$ B signaling, a crucial pathway in mediating endothelial activation.

Congestive heart failure (CHF) is a CVD that develops in part as a complication of atherosclerosis that is characterized by the inability of the heart to pump blood. Using a novel genetic screening panel in mouse in combination with system based approaches, I have identified and characterized the function of numerous novel genes that modulate CHF phenotypes such as cardiac fibrosis (*Abcc6*), left ventricular mass (*Myh14*) and cardiac hypertrophy (*Adamts2*).

The dissertation of Milagros De La Caridad Romay is approved.

Sherie L. Morrison

Yibin Wang

Jerome A. Zack

Aldons J. Lulis

University of California, Los Angeles

2016

## **DEDICATION**

My dissertation is dedicated to my mother. She who is my inspiration, my guiding force, and my companion during the last ten years of my education. This is one more hurdle on the long road to success that I have overcome with your help. I thank you for everything and look forward to better days.

And I dedicate this dissertation to my teachers:

Dr. Mete Civelek

Dr. Judith A. Berliner

Dr. Jake Lusi

Thank you for showing me the way and for inspiring me to be not only a better scientist, but a better person.

## TABLE OF CONTENTS

Abstract of the Dissertation	II
Dedication	V
List of Figures	VIII
List of Tables	XI
Acknowledgements	XIII
Vita	XVI
Publications	XVII
<b>Chapter 1: Introduction.</b>	<b>1</b>
References	<b>9</b>
<b>Chapter 2: Regulation of NF-<math>\kappa</math>B signaling by oxidized glycerophospholipid and IL-1<math>\beta</math> induced miRs-21-3p and -27a-5p in human aortic endothelial cells.</b>	<b>12</b>
References	<b>24</b>
<b>Chapter 3: Mapping genetic contributions to cardiac pathology induced by Beta-adrenergic stimulation in mice.</b>	<b>27</b>
References	<b>37</b>
<b>Chapter 4: Genetic Dissection of Cardiac Remodeling in an Isoproterenol-induced Heart Failure Mouse Model.</b>	<b>39</b>
References	<b>91</b>
<b>Chapter 5: A systems genetics approach to identify genetic pathways and key drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in mice.</b>	<b>99</b>

References	152
<b>Chapter 6: Concluding Remarks.</b>	<b>163</b>
References	173

## LIST OF FIGURES

### CHAPTER 2

- Figure 2-1:** Highly expressed HAEC miRNAs. 17
- Figure 2-2:** Expression of select miRNAs in HAECs in response to various stimuli. 18
- Figure 2-3:** miR-21-3p and miR-27a-5p inhibit NF- $\kappa$ B signaling in HAECs. 20
- Figure 2-4:** Identification of candidate target genes of miRs-21-3p and -27a-5p. 21
- Figure 2-5:** miRs-21-3p and -27a-5p affect the expression of multiple genes in the NF- $\kappa$ B signaling cascade. 22

### CHAPTER 3

- Figure 3-1:** Wide variation in heart failure (HF) traits between the strains of the hybrid mouse diversity panel (HMDP). 33
- Figure 3-2:** Manhattan plots of heart failure (HF) traits. 35
- Figure 3-3:** *Ppp3ca* (calcineurin A) is a candidate gene at the Chromosome 3 right ventricular weight ratio locus. 36
- Figure 3-4:** *Abcc6* plays a role in the regulation of cardiac fibrosis after isoproterenol (ISO) stimulation. 37

### CHAPTER 4

- Figure 4-1:** Experimental Design. 82

<b>Figure 4-2:</b> Variation in echocardiographic measures of cardiac structure and function among HMDP mouse strains.	<b>83</b>
<b>Figure 4-3:</b> Population distribution of echocardiographic measures at each time point.	<b>84</b>
<b>Figure 4-4:</b> Correlations among echocardiographic traits across time points.	<b>85</b>
<b>Figure 4-5:</b> Principal component analysis of isoproterenol-induced changes from baseline.	<b>86</b>
<b>Figure 4-6:</b> Genome-wide association for week 3 LVM.	<b>87</b>
<b>Figure 4-7:</b> Regional plots of LV hypertrophy at chromosome 7 across time points.	<b>88</b>
<b>Figure 4-8:</b> <i>Myh14</i> knock-down in stressed neonatal rat ventricular myocytes (NRVMs) causes abrogation of hypertrophic response and decreased cell viability.	<b>89</b>
<b>Figure 4-9:</b> <i>Myh14</i> deficiency under ISO results in LVM hypertrophy, LVIDd dilation, increased cardiomyocyte size, fibrosis, intercalated disc disarray, and hypertrophic signals.	<b>90</b>
<b>Figure 4-10:</b> Relationships between variation and correlated traits.	<b>91</b>

## CHAPTER 5

<b>Figure 5-1:</b> Visualization and functional enrichment of the left ventricular gene co-expression network.	<b>131</b>
<b>Figure 5-2:</b> Module - HF Trait Correlation.	<b>132</b>
<b>Figure 5-3:</b> Visualization and functional enrichment of Module 5.	<b>133</b>
<b>Figure 5-4:</b> Near Edge Orientation Algorithm.	<b>134</b>

<b>Figure 5-5:</b> <i>Adamts2</i> alters the expression of module 5 genes in cardiomyocytes.	<b>135</b>
<b>Figure 5-6:</b> <i>Adamts2</i> regulates cardiac hypertrophy <i>in vitro</i> .	<b>136</b>
<b>Supplemental Figure 5-1:</b> Module-Trait Correlation for Control Network.	<b>146</b>
<b>Supplemental Figure 5-2:</b> Module-Trait Correlation for Treated Network.	<b>147</b>
<b>Supplemental Figure 5-3:</b> Module-Trait Correlation for Induction Network.	<b>148</b>
<b>Supplemental Figure 5-4:</b> Graphical depiction of module 1 in the control network.	<b>149</b>
<b>Supplemental Figure 5-5:</b> Near Edge Orientation of Module 19 of the treated network.	<b>150</b>
<b>Supplemental Figure 5-6:</b> Near Edge Orientation of Module 14 of the delta network.	<b>151</b>
<b>Supplemental Figure 5-7:</b> Hierarchical Clustering of Induction Module Eigengenes.	<b>152</b>

## LIST OF TABLES

### CHAPTER 2

<b>Table 2-1:</b> Abundance of miR-21 and miR-27a (percentage) in HAECs.	<b>17</b>
<b>Table 2-2:</b> Functional enrichment of Gene Ontology categories of genes downregulated by miR-21-3p and miR-27a-5p.	<b>18</b>
<b>Table 2-3:</b> Putative gene targets of miR-21-3p and miR-27a-5p for validation.	<b>20</b>

### CHAPTER 3

<b>Table 3-1:</b> Significant Heart Failure Trait Loci Identified in HMDP GWA.	<b>33</b>
<b>Table 3-2:</b> Suggestive Heart Failure Trait Loci Identified in HMDP GWA.	<b>34</b>

### CHAPTER 4

<b>Table 4-1:</b> Heritability estimates of cardiac structure and function under isoproterenol treatment.	<b>76</b>
<b>Table 4-2:</b> Significant association loci.	<b>77</b>

### CHAPTER 5

<b>Table 5-1:</b> Top Gene Ontology Enrichments in Modules of the left ventricular Gene Coexpression Network.	<b>130</b>
<b>Supplementary Table 5-1:</b> Strains used in this study.	<b>137</b>
<b>Supplemental Table 5-2:</b> Top Gene Ontologies for Modules of the Control Left Ventricular MICA Network.	<b>140</b>
<b>Supplemental Table 5-3:</b> Genes in Module 1 of the treated network with significant correlation to RV weight.	<b>141</b>

<b>Supplemental Table 5-4:</b> Top Gene Ontologies for Modules of the Ratio of Treated to Control Left Ventricular MICA Network.	<b>142</b>
<b>Supplemental Table 5-5:</b> Drivers and reactive genes in module 5.	<b>143</b>
<b>Supplemental Table 5-6:</b> Characteristics of Module 5 <i>Adamts2</i> target genes that were selected for <i>in vitro</i> validation using siRNA knockdown in NRVMs.	<b>145</b>

## ACKNOWLEDGEMENTS

I would like to acknowledge Jake Lusic for introducing me to genetics during the fall quarter of 2008 and spring quarter of 2009, and again for giving me a chance to work with him first as a master's student in 2010, and again as a PhD student in 2012. Truly one of a kind, his enthusiasm for science and dedication for his trainees has been a source of inspiration during my time in his lab. There are not enough words to say how thankful I am to have been able to train under Jake both during my master's and PhD. It is my time in his lab that has truly transformed me not only as a scientist but as a human being.

I would also like to acknowledge Judith Berliner for her dedication and encouragement during my graduate work. During my training she was a constant presence in the lab, a true touchstone of wisdom and guidance with a no-nonsense attitude that I hope one day to emulate with my own trainees. She is a constant source of support and an excellent role model and I am blessed to have her support during even now as I continue on to my post-doctoral training.

In addition, I would like to acknowledge a final mentor in the Lusic Lab, Mete Civelek. I cannot thank him enough for the time and energy he spent in training me during his time in the Lusic lab as a postdoctoral fellow. In particular, it was his mentorship and guidance that played a crucial role during my training in 2010 through 2014 which cumulated in the manuscript included in this dissertation as Chapter 2 "Regulation of NF- $\kappa$ B signaling by oxidized phospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells".

They say it takes a village to raise a child, but it is only looking back at my training that I realize it takes an entire lab to train a PhD student. I would like to acknowledge the members of the Lusic Lab for their support and advice during my training. I would specifically like to

acknowledge my collaborators Christoph Rau and Jessica Wang, who were major forces behind the body of work included in this dissertation and responsible for generating the Heart Failure HMDP data which serves as the basis for Chapters 3,4 and 5 of this dissertation. In addition, I would like to acknowledge a few of the individual contributions of specific Lusis lab members during my training: Brian Parks, for advice and guidance during the early years of my PhD and Marcus Seldin, for advice and guidance during the later years of my training. I would like to also thank Nam Che for technical support and advice and Rosa Chen for administrative support and advice during my six years in the Lusis Lab.

In Chapter 2, I have included a reproduction of the manuscript “Regulation of NF- $\kappa$ B signaling by oxidized phospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells”. Romay, M.C., et al., *Journal of Lipid Research* 2015 Jan;56(1)p:38-50. Per the copyright permission policy of the *Journal of Lipid Research*, published by the American Society for Biochemistry and Molecular Biology, I have permission to reuse this manuscript as part of a dissertation.

In Chapter 3, I have included a reproduction of the manuscript “Mapping Genetic Contributions to Cardiac Pathology Induced by Beta-Adrenergic Stimulation in Mice”. Rau, C.D., et al. *Circulation Cardiovascular Genetics* 2015 Feb;8(1):40-9. Per Permissions Policy of *Circulation: Cardiovascular Genetics* an AHA Journal published by Wolters Kluwer Health, Lippincott Williams & Wilkins, I have permission as an author to reuse this manuscript as part of a dissertation.

Chapter 4 is a version of a submitted manuscript “Genetic Dissection of Cardiac Remodeling in an Isoproterenol-induced Heart Failure Mouse Model” Jessica Jen-Chu Wang,

Christoph D. Rau, Rozeta Avetisyan, Shuxun Ren, Milagros C. Romay, Gabriel Stolin, Kei Wei Gong, Yibin Wang, Aldons J. Lusic. JJW, CDR, YW and AJL designed the research, JJW, CDR, RA, SR, MCR, GS, KWG performed research and designed figures, JJW wrote the manuscript.

Chapter 5 is a version of a submitted manuscript “A systems genetics approach to identify genetic pathways and key drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in mice.” Christoph D. Rau, Milagros C. Romay, Jessica J-C. Wang, Shuxun Ren, Yibin Wang, Aldons J. Lusic. CDR, MCR, JJW, YW and AJL designed the research, CDR, MCR and JJW performed the research, CDR and MCR wrote the manuscript.

I received funding from University of California Office of the President and the UCLA Graduate Division as a recipient of the Eugene V. Cota Robles Fellowship during the 2011-2012, and 2014-2015 academic years in addition to support from the UCLA Cellular and Molecular Biology training grant from 2013-2014 and 2015-2016 academic years (Ruth L. Kirschstein National Research Service Award, Grant T32GM718539).

## VITA

- 2010 B.S., Microbiology, Immunology and Molecular Genetics  
Minor: Women's Studies, Society and Genetics  
University of California, Los Angeles
- 2010-Present Graduate Student Research  
Department of Microbiology, Immunology and Molecular  
Genetics  
University of California, Los Angeles.
- 2011-2012 University of California Office of the President and UCLA  
Graduate Division - Eugene V. Cota Robles Fellowship Year  
1
- 2013 Teaching Assistant, Human Genetics 156  
Department of Microbiology, Immunology and Molecular  
Genetics,  
University of California, Los Angeles
- 2013-2014 NIH Ruth L. Kirschstein National Research Service Award  
"Identification and Characterization of miR-30d as a Novel  
Regulator of Cardiac Fibrosis"
- 2014-2015 University of California Office of the President and UCLA  
Graduate Division - Eugene V. Cota Robles Fellowship Year  
2
- 2015 Teaching Assistant, Human Genetics 156  
Department of Microbiology, Immunology and Molecular  
Genetics,  
University of California, Los Angeles
- 2015-2016 NIH Ruth L. Kirschstein National Research Service Award  
"*Adams2* as a novel regulator of cardiac hypertrophy"

#### Publications:

1. Romay MC, Che N, Becker SN, Pouldar D, Hagopian R, Xiao X, Lusic AJ, Berliner JA, Civelek M. Regulation of NF- $\kappa$ B signaling by oxidized glycerophospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells. *J. Lipid Research*
2. Rau CD, Wang J, Avetisyan R, Romay MC, Martin L, Ren S, Wang Y, Lusic AJ. Mapping Genetic Contributions to Cardiac Pathology Induced by Beta-adrenergic Stimulation in Mice. *Circ. Caridovasc. Genet.*

#### Posters:

1. Romay MC, Che N, Becker SN, Pouldar D, Hagopian R, Xiao X, Lusic AJ, Berliner JA, Civelek M. Regulation of NF- $\kappa$ B signaling in Human Aortic Endothelial Cells by Oxidized Phospholipid induced microRNAs. NAVBO Vascular Biology 2012.
2. Rau CD, Avetisyan R, Romay M, Stein D, Wang J, Wang Y, Lusic AJ. Genetic Basis of Isoproterenol-induced Cardiac Fibrosis. American Heart Association Basic Cardiovascular Science Meeting 2014.

#### Invited Talks

3. Regulation of NF- $\kappa$ B signaling by oxidized phospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells. NAVBO Vascular Biology 2014

## **Chapter 1**

### **Introduction**

The human body is a large and complex biological unit composed of interconnecting organ systems formed from cells and tissues that communicate with each other to perform the necessary functions for life. The principal function of the cardiovascular system is to regulate the transport of oxygen and nutrients to tissues while removing excess carbon dioxide and waste from the periphery. To perform this function the cardiovascular system is composed of two primary components: the heart which acts to provide the necessary force to pump blood through the vasculature allowing for the transport solutes and oxygen to the periphery and the vasculature, the site of exchange for nutrients and waste with the periphery.

Over the last 100 years, cardiovascular diseases have overtaken infectious agents as the primary cause of mortality across the globe, representing 30% of all deaths that occur globally each year<sup>1</sup>. Despite the success of focused, reductionist approaches in characterizing interactions that contribute to these diseases, current estimates predict that 23.6 million deaths annually will be due to CVDs by 2030<sup>1</sup>. Systems genetics is a holistic, integrative approach to understanding how the molecular interactions in a biological system on a global scale contribute to the observed biological phenomena in question utilizing genetic variation as the key perturbation<sup>2</sup>. Characterized by the integration next generation –omics technologies, bioinformatics and molecular biology techniques, systems genetics allow investigators to integrate data across many levels to identify key regulators in biological processes. As common CVDs involve nonlinear genetic and environmental interactions that are difficult to dissect using traditional approaches, systems genetics methodologies offer an appealing alternative for the identification of novel genes contributing to these disorders<sup>3</sup>.

### **Regulation of gene expression during endothelial activation by microRNAs**

Atherosclerosis, a systemic disorder characterized by the narrowing of arteries is the underlying cause of the majority of clinical cardiovascular events including myocardial infarction and stroke<sup>4</sup>. Formation of the atherosclerotic plaque is driven by chronic endothelial activation from exposure to oxidized phospholipids and cytokines that accumulate within the vessel wall<sup>5</sup>. Exposure of endothelial cells to oxidized lipids such as oxidized 1-palmitoyl-2-arachidonoyl-sn-glycero-3-phosphorylcholine (OxPAPC) which accumulates in atherosclerotic lesions, perturbs the expression of thousands of genes involved with vital endothelial cell biological process including unfolded protein response, inflammation and cell cycle<sup>6</sup>. Regulation of gene expression on this scale can occur through multiple molecular processes including induction of microRNAs. MicroRNAs (miRNAs) are short non-coding RNAs that modulate gene expression through translational repression and degradation of their messenger RNA<sup>7</sup>.

In chapter 2 “Regulation of NF- $\kappa$ B signaling by oxidized glycerophospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells” I describe the identification of two novel microRNAs, miRs-21-3p and miR-27a-5p<sup>8</sup>. Using next generation sequencing in combination with classical molecular biology approaches I have identified miRs-21-3p and -27a-5p as novel regulators of NF- $\kappa$ B signaling in response to inflammatory stimuli in endothelial cells. I show in the course of this chapter that miRs-21-3p and -27a-5p act to regulate the NF- $\kappa$ B cascade at multiple points in the cascade in addition to fine-tuning activation of the cascade through modulation of both activators and regulators of NF- $\kappa$ B signaling in endothelial cells. As NF- $\kappa$ B signaling is crucial in mediating endothelial activation the data presented in this chapter strongly suggest a role for miRs-21-3p and -27a-5p in maintaining the tight regulation of NF- $\kappa$ B pathway to maintain proper vessel function

## **From Atherosclerosis to Heart Failure**

Rupture of the atherosclerotic lesion in the coronary vessels of the heart is the main cause for myocardial infarction, the irreversible death of the myocardium due to lack of oxygen. While advances in medical treatment have caused a decreased in mortality due to MI, many patients who do survive are predicted to develop secondary complications from this event including congestive heart failure. Congestive heart failure (CHF) is a CVD that is characterized by the inability of the heart to pump blood due excessive remodeling following stress leading to the development of peripheral edema, shortness of breath, reduced quality of life and eventual death<sup>9</sup>. CHF is an enormous burden on the public health system as it is the leading cause of hospitalization among the elderly in the US with total medical costs for patients suffering from CHF expected to \$53.1 billion by 2030<sup>10</sup>.

While MI is one of the leading risk factors for development of CHF, numerous other environmental and genetic risk factors exist including hypertension, smoking, obesity and diabetes. Despite the success of population based screening approaches in identifying genetic components in other CVDs such as atherosclerosis, atrial fibrillation and hypertension attempts to utilize these techniques in heart failure have shown limited success due to the disease's complex etiology<sup>11-13</sup>.

## **Systems Genetics of Heart Failure in the HMDP**

The hybrid mouse diversity panel (HMDP) is a unique mouse reference genetic population developed in the Lusis lab to perform systems genetics analysis of complex traits. The HMDP is a combination over 100 classical and recombinant inbred (RI) mouse strains<sup>14</sup>. Over the last five years utilization of the HMDP for system genetics analysis has identified novel

genes and biological pathways in such complex traits as atherosclerosis, bone mineral density, non-alcoholic fatty liver disease, immune response, obesity and insulin resistance<sup>15-21</sup>.

To model the development of CHF in the HMDP, HMDP mice were submitted to a 21-day isoproterenol treatment regimen. Isoproterenol, which models the chronic  $\beta$ -adrenergic stimulation that often occurs in human heart failure, was administered by an implanted pump for 3 weeks. During this three week period echocardiographic measures were collected and following the 21 day treatment period the mice were sacrificed to collect tissue and histological measurements. During the course of the heart failure HMDP (HF HMDP), a total of 734 mice and 48 quantitative traits were collected and analyzed.

### **Systems Genetics approaches in Heart Failure – Genome Wide Association Studies Identify Novel Loci for Cardiac Fibrosis and Left Ventricular Mass.**

One of the main tools in the system geneticists' repertoire for the identification of novel genes contributing to complex traits is gene mapping. In the mouse genetics, analysis of complex traits using traditional F2 crosses generate broad genomic regions related to the trait due to recombination frequency. In comparison, the HMDP's combination of classical inbred strains and 4 RI panels, provide enhanced resolution and power to map complex traits using association generating narrower genomic regions of significant association for further investigation<sup>14</sup>.

In Chapter 3 "Mapping genetic contributions to cardiac pathology induced by Beta-adrenergic stimulation in mice." I describe the results of the genome wide association study (GWAS) for the HF HMDP organ weight traits and cardiac fibrosis<sup>22</sup>. Our analyses revealed 7 significant loci and 17 suggestive loci for traits related to the development of cardiac hypertrophy and fibrosis. In particular, we identified the gene, *Abcc6* as a novel contributing factor to the development of cardiac fibrosis locus by employing both knockout and transgenic

mouse models. Cardiac fibrosis, the accumulation of fibrotic tissue within the myocardium that impairs both mechanical and conductive properties of the myocardium is an important risk factor for CHF that is poorly understood. Identification of *Abcc6* as a novel risk factor for the development of cardiac fibrosis emphasizes the utility of the HMDP for identifying genetic factors for complex traits that are difficult to characterize in humans.

In Chapter 4 “Genetic Dissection of Cardiac Remodeling in an Isoproterenol-induced Heart Failure Mouse Model” I describe the results of the GWAS for the HF HMDP echocardiographic traits. Serial echocardiography is a powerful and noninvasive tool to serially monitor changes in cardiac structure and function in humans in response to stress and injury. Using this approach we successfully identified 17 genome-wide significant loci associated with indices of cardiac remodeling. Furthermore, we showed that genetic variation in a novel gene *Myh14* affects heart failure by altering the mechanical responses of heart muscles to isoproterenol-induced stress using a combination of in vitro and in vivo models.

### **Systems Genetics Applications in the HMDP – Gene Co-expression Network Modeling**

The integration of genetic information with gene expression network modeling has been successfully applied to identify candidate genes that regulate biological processes in other complex diseases such as coronary artery disease, and Alzheimer's disease<sup>2</sup>. In Chapter 5 “A systems genetics approach to identify genetic pathways and key drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in mice.” I describe the construction and characterization of a gene co-expression network of left ventricular transcriptomes (LV) from 92 HMDP strains pre- and post- 21 day isoproterenol treatment applying the non-linear network analysis algorithm, MICA. This ISO-LV expression network was composed of 8,126 genes, and contained 20 modules, each containing a variable number of highly correlated genes. Correlation

of the module eigengenes, identified the 5th module (termed module 5) as significantly correlated with seven heart failure traits including left ventricular weight (a proxy trait for cardiac hypertrophy).

Applying the causality modeling algorithm NEO to module 5, we identified the gene *Adamts2* as a putative master regulator of this module. Using siRNA knockdown of *Adamts2* in cardiomyocytes we identified 5 new targets of *Adamts2*; *Kcnv2*, *Mfap2*, *Tnc*, *Nppa* and *Nppb*. Of these, *Tnc*, *Nppa* and *Nppb* are well associated with the development of heart failure, while *Kcnv2* and *Mfap2* represent two novel genes not previously implicated in the development of heart failure. The results suggested that *Adamts2*, a metalloproteinase not previously associated with HF, plays a key role in driving the reaction of other genes in the module in response to ISO stimulation.

### **Systems Genetics Framework for Analysis of CVDs**

Despite public health efforts and improved patient care cardiovascular diseases such as atherosclerosis and congestive heart failure are growing causes of morbidity and mortality worldwide. With the decreasing costs of next generation sequencing technologies, the ability of scientist to investigate multiple scales of biological data and integrate that data to understand the broad molecular underpinnings of disease is closer to becoming a reality. Regardless, their still remains numerous challenges both in designing and implementing system genetics studies and analyzing the massive datasets generated from these studies to gain novel insights in to complex traits such as CVDs.

In summary, my dissertation work addresses questions involving the applicability of the systems genetics in both complex and inherited CVDs. In reference to atherosclerosis, I have identified and characterized two novel microRNAs involved in the response of endothelial cells

to inflammatory stimuli, a crucial process in the development of the atherosclerotic lesions. In relation to congestive heart failure, using systems genetics approaches I have identified and characterized three novel genes contributing to the development of cardiac fibrosis (*Abcc6*), left ventricular mass as determined by echocardiography (*Myh14*) and cardiac hypertrophy (*Adamts2*). Finally, I've applied the systems genetics approaches I have learned during the course of the atherosclerosis and congestive heart failure projects to a novel application – analysis of exome sequencing data from a familial case of arrhythmogenic right ventricular cardiomyopathy.

## References

- 1 Wong, N. D. Epidemiological studies of CHD and the evolution of preventive cardiology. *Nat Rev Cardiol* **11**, 276-289, doi:10.1038/nrcardio.2014.26 (2014).
- 2 Civelek, M. & Lusis, A. J. Systems genetics approaches to understand complex traits. *Nat Rev Genet* **15**, 34-48, doi:10.1038/nrg3575 (2014).
- 3 MacLellan, W. R., Wang, Y. & Lusis, A. J. Systems-based approaches to cardiovascular disease. *Nat Rev Cardiol* **9**, 172-184, doi:10.1038/nrcardio.2011.208 (2012).
- 4 Lusis, A. J. Atherosclerosis. *Nature* **407**, 233-241, doi:10.1038/35025203 (2000).
- 5 Libby, P. Inflammation in atherosclerosis. *Nature* **420**, 868-874, doi:10.1038/nature01323 (2002).
- 6 Romanoski, C. E. *et al.* Network for activation of human endothelial cells by oxidized phospholipids: a critical role of heme oxygenase 1. *Circ Res* **109**, e27-41, doi:10.1161/CIRCRESAHA.111.241869 (2011).
- 7 Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233, doi:10.1016/j.cell.2009.01.002 (2009).
- 8 Romay, M. C. *et al.* Regulation of NF-kappaB signaling by oxidized glycerophospholipid and IL-1beta induced miRs-21-3p and -27a-5p in human aortic endothelial cells. *J Lipid Res* **56**, 38-50, doi:10.1194/jlr.M052670 (2015).
- 9 Mudd, J. O. & Kass, D. A. Tackling heart failure in the twenty-first century. *Nature* **451**, 919-928, doi:10.1038/nature06798 (2008).
- 10 Ziaeian, B. & Fonarow, G. C. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol*, doi:10.1038/nrcardio.2016.25 (2016).

- 11 Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121-1130, doi:10.1038/ng.3396 (2015).
- 12 Wain, L. V. *et al.* Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet* **43**, 1005-1011, doi:10.1038/ng.922 (2011).
- 13 Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat Genet* **44**, 670-675, doi:10.1038/ng.2261 (2012).
- 14 Ghazalpour, A. *et al.* Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm Genome* **23**, 680-692, doi:10.1007/s00335-012-9411-5 (2012).
- 15 Bennett, B. J. *et al.* A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Research* **20**, 281-290, doi:10.1101/gr.099234.109 (2010).
- 16 Bennett, B. J. *et al.* Genetic Architecture of Atherosclerosis in Mice: A Systems Genetics Analysis of Common Inbred Strains. *PLOS Genetics* **11**, e1005711, doi:10.1371/journal.pgen.1005711 (2015).
- 17 Farber, C. R. *et al.* Mouse Genome-Wide Association and Systems Genetics Identify *Asxl2* As a Regulator of Bone Mineral Density and Osteoclastogenesis. *PLoS Genetics* **7**, e1002038, doi:10.1371/journal.pgen.1002038 (2011).
- 18 Hui, S. T. *et al.* The genetic architecture of NAFLD among inbred strains of mice. *Elife* **4**, e05607, doi:10.7554/eLife.05607 (2015).

- 19 Orozco, Luz D. *et al.* Unraveling Inflammatory Responses using Systems Genetics and Gene-Environment Interactions in Macrophages. *Cell* **151**, 658-670, doi:10.1016/j.cell.2012.08.043 (2012).
- 20 Parks, Brian W. *et al.* Genetic Control of Obesity and Gut Microbiota Composition in Response to High-Fat, High-Sucrose Diet in Mice. *Cell Metabolism* **17**, 141-152, doi:10.1016/j.cmet.2012.12.007 (2013).
- 21 Parks, Brian W. *et al.* Genetic Architecture of Insulin Resistance in the Mouse. *Cell Metabolism* **21**, 334-346, doi:10.1016/j.cmet.2015.01.002 (2015).
- 22 Rau, C. D. *et al.* Mapping genetic contributions to cardiac pathology induced by Beta-adrenergic stimulation in mice. *Circ Cardiovasc Genet* **8**, 40-49, doi:10.1161/CIRCGENETICS.113.000732 (2015).

## Chapter 2

**Regulation of NF- $\kappa$ B signaling by oxidized glycerophospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells.**

**Chapter 2 Preface:** This chapter is a reprint of a manuscript describing the identification and characterization of two novel microRNAs, miRs-21-3p and -27a-5p in human aortic endothelial cells. Utilizing a combination of approaches including next generation sequencing, whole genome transcript profiling using microarrays, in-vitro microRNA overexpression and knockdown, RT-qPCR, western blot and 3'UTR luciferase assay I identified miRs-21-3p and -27a-5p as modulators of NF- $\kappa$ B signaling in endothelial cells upon exposure to oxidized phospholipids and IL-1 $\beta$ . As first author on this publish manuscript I was responsible for the experimental design to identify the biological processes regulated by miRs-21-3p and -27a-5p, I preformed all experiments with the exception of the HAEC miRNA deep sequencing, I processed and analyzed all the data with the exception of the deep sequencing mapping and normalization of the miR-21-3p and 27a-5p overexpression microarrays, generated all the figures and was the major contributor to the manuscript text.

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

## Regulation of NF- $\kappa$ B signaling by oxidized glycerophospholipid and IL-1 $\beta$ induced miRs-21-3p and -27a-5p in human aortic endothelial cells<sup>[S]</sup>

Milagros C. Romay,\* Nam Che,<sup>†</sup> Scott N. Becker,\* Delila Pouldar,<sup>†</sup> Raffi Hagopian,<sup>§</sup> Xinshu Xiao,\*\* Aldons J. Lusis,\*<sup>†,§</sup> Judith A. Berliner,<sup>††</sup> and Mete Civelek<sup>1,†</sup>

Departments of Microbiology, Immunology and Molecular Genetics,\* Medicine,<sup>†</sup> Human Genetics,<sup>§</sup> Integrative Biology and Physiology,\*\* and Pathology and Laboratory Medicine,<sup>††</sup> University of California, Los Angeles, Los Angeles, CA 90095

**Abstract** Exposure of endothelial cells (ECs) to agents such as oxidized glycerophospholipids (oxGPs) and cytokines, known to accumulate in atherosclerotic lesions, perturbs the expression of hundreds of genes in ECs involved in inflammatory and other biological processes. We hypothesized that microRNAs (miRNAs) are involved in regulating the inflammatory response in human aortic endothelial cells (HAECs) in response to oxGPs and interleukin 1 $\beta$  (IL-1 $\beta$ ). Using next-generation sequencing and RT-quantitative PCR, we characterized the profile of expressed miRNAs in HAECs pre- and postexposure to oxGPs. Using this data, we identified miR-21-3p and miR-27a-5p to be induced 3- to 4-fold in response to oxGP and IL-1 $\beta$  treatment compared with control treatment. Transient overexpression of miR-21-3p and miR-27a-5p resulted in the downregulation of 1,253 genes with 922 genes overlapping between the two miRNAs. Gene Ontology functional enrichment analysis predicted that the two miRNAs were involved in the regulation of nuclear factor  $\kappa$ B (NF- $\kappa$ B) signaling. Overexpression of these two miRNAs leads to changes in p65 nuclear translocation. Using 3' untranslated region luciferase assay, we identified 20 genes within the NF- $\kappa$ B signaling cascade as putative targets of miRs-21-3p and -27a-5p, implicating these two miRNAs as modulators of NF- $\kappa$ B signaling in ECs.—Romay, M. C., N. Che, S. N. Becker, D. Pouldar, R. Hagopian, X. Xiao, A. J. Lusis, J. A. Berliner, and M. Civelek. Regulation of NF- $\kappa$ B signaling by oxidized glycerophospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells. *J. Lipid Res.* 2015. 56: 38–50. **Supplementary key words** cell signaling • cytokines • gene expression • inflammation oxidized lipids • microRNAs • nuclear factor  $\kappa$ B • interleukin 1 $\beta$

This work was supported by National Institutes of Health Grants HL-30568 and HL-28481 (A.J.L. and J.A.B.); Ruth L. Kirschstein National Research Service Award, Grant T32GM718539 (M.C.R.); American Heart Association postdoctoral fellowship, Grant 10POST3660048; National Institutes of Health Ruth L. Kirschstein National Research Service Award, Grant T32HL69766; and National Institutes of Health Pathway to Independence Award, Grant K99HL121172 (M.C.).

Manuscript received 1 July 2014 and in revised form 26 September 2014.

Published, JLR Papers in Press, October 19, 2014

DOI 10.1194/jlr.M052670

The endothelium transitions from a quiescent state to an active state during the initiation of the innate immune response. The active endothelium is characterized by an increased expression of leukocyte adhesion molecules and chemotactic factors as well as increased permeability of the vascular beds allowing for leukocyte recruitment into the periphery from the blood. The persistence of endothelial activation due to chronic exposure of inflammatory stimuli is critical to the development of many diseases, including atherosclerosis (1).

During the initiation of atherosclerosis, oxidized low density lipoprotein (oxLDL) trapped in the vessel wall contributes to the activation of ECs (2). In vitro treatment of human aortic endothelial cells (HAECs) with a component of oxLDL, oxidized 1-palmitoyl-2-arachidonoyl-*sn*-glycero-3-phosphatidylcholine (Ox-PAPC), changes the expression of hundreds of genes involved in inflammation, unfolded protein response, coagulation, and sterol biosynthesis (3–8). Similar global effects of interleukin 1 $\beta$  (IL-1 $\beta$ ) and TNF $\alpha$  on ECs have been observed (9). However, the regulation of the response to these stimuli has not been fully characterized. We hypothesized that miRNAs partially regulate the expression of genes that are involved in the response of ECs to Ox-PAPC.

MicroRNAs (miRNAs) are small noncoding RNAs that can affect the expression of hundreds to thousands of

Abbreviations: Ago, Argonaute; AKT1, v-akt murine thymoma viral oncogene homolog 1; AP-1, activator protein 1; EC, endothelial cell; GM-CSF, granulocyte macrophage-colony-stimulating factor; HAEC, human aortic endothelial cell; HUVEC, human umbilical vein endothelial cell; IL-1 $\beta$ , interleukin 1 $\beta$ ; miRNA, microRNA; NF- $\kappa$ B, nuclear factor  $\kappa$ B; NK, natural killer; oxGP, oxidized glycerophospholipid; Ox-PAPC, oxidized 1-palmitoyl-2-arachidonoyl-*sn*-glycero-3-phosphatidylcholine; UTR, untranslated region; VCAM-1, vascular cell adhesion molecule-1.

<sup>1</sup>To whom correspondence should be addressed.

e-mail: [mcivelek@mednet.ucla.edu](mailto:mcivelek@mednet.ucla.edu)

<sup>[S]</sup> The online version of this article (available at <http://www.jlr.org>) contains supplementary data in the form of eight figures, two tables, and a supplementary file.

Copyright © 2015 by the American Society for Biochemistry and Molecular Biology, Inc.

This article is available online at <http://www.jlr.org>

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

genes by binding to their target mRNAs causing translational repression or target degradation (10). miRNAs contribute to the regulation of many endothelial functions including angiogenesis, vasodilation, coagulation, and inflammation (11–14). Furthermore, expression of EC miRNAs is highly sensitive to environmental cues. Common risk factors for atherosclerosis including shear stress, high blood glucose, and oxidized lipids have been shown to alter EC miRNA expression suggesting a role for miRNAs during disease pathology (15–18).

In this study, we first characterized the repertoire of miRNAs expressed in HAECs using next-generation sequencing of small RNAs. We then identified miRs-21-3p and -27a-5p to be responsive to Ox-PAPC, IL-1 $\beta$  and TNFC. We found that these two miRNAs act to regulate hundreds of genes, including genes associated with nuclear factor KB (NF-KB) signaling in ECs. We demonstrate that transient overexpression of miRs-21-3p and -27a-5p leads to decreased p65 nuclear translocation in response to proinflammatory stimulant IL-1 $\beta$  suggesting that miRs-21-3p and -27a-5p act to repress NF-KB signaling in HAECs. We show using 3' untranslated region (UTR) luciferase assay that miRs-21-3p and -27a-5p alter the expression of both repressors and activators of the NF-KB signaling cascade suggesting that these two miRNAs act to regulate the extent of NF-KB signaling in ECs.

## METHODS

### Cell culture and treatments

HAECs were isolated from aortas of human heart transplant donors, and cultures were maintained in MCDB-131C complete medium (VEC Technologies, Rensselaer, NY). Ox-PAPC was prepared as described previously (19). For some studies, cells were treated in 1% FBS medium containing 40  $\mu$ g/ml Ox-PAPC or vehicle for 4 h. In other studies, cells were treated with 20 ng/ml or 2 ng/ml IL-1 $\beta$  (201-LB-005/CF) and 20 ng/ml TNFC (210-TA) purchased from R and D Systems in 1% FBS medium for 2 h.

miRNA overexpression experiments were performed using Dharmacon miRDIAN miRNA mimics (miR-21-3p: C-301023-01-0005; miR-27a-5p: C-301028-01-0005; negative control: CN-001000-01-05) at indicated concentrations for 24 h. miRNA knockdown experiments were performed using the Exiqon miRCURY LNA<sup>TM</sup> miRNA inhibitors (miR-21-3p: 410136-00; miR-27a-5p: 410168-00; negative control: 199004-00) at the indicated concentrations for 24 h. All transfections were performed with Invitrogen Lipofectamine 2000 reagent.

### miRNA library preparation and sequencing

Following treatment, cells were lysed, and total RNA was isolated using the Qiagen miRNeasy Kit to retain the small RNA fraction. RNA integrity number (RIN) values were assessed with the Agilent Bioanalyzer 2100 instrument. Samples with RIN values > 9.0 were used for transcriptional profiling. Small RNA libraries were prepared using the Illumina Small RNA v1.5 protocol and sequenced for 76 base reads on the Illumina GAIIx platform.

### Alignment and quantification of miRNA deep sequencing

The small RNA sequencing data analysis method has been described in detail elsewhere (20). Briefly, the sequencing files for

the four HAEC libraries were converted to the FASTQ format using a custom Perl script. The reads were then aligned to the hg19 version of the genome using the Novoalign tool with the following settings: -116 -t30 -h90 -rA -R 1 -m -g 200 -k. These settings allowed for a single read to map to multiple regions of the genome with up to one mismatch. We quantified the number of reads aligning to the genomic coordinates of known mature miRNAs downloaded from miRBase version 19 using the Bioconductor package GenomicRanges for R (v.2.14.0) (21). Whenever a read mapped to "x" genomic loci, the read would contribute a count of 1/x to those regions. To enable comparison of counts between samples, we normalized the expression values by dividing the counts for a given mature miRNA by the sum of all the miRNA counts for the corresponding library.

### Whole genome transcript profiling and differential gene expression analysis

HAECs from two different donors were transfected with 1 nM miRNA mimics and negative control for 24 h in triplicate. Total RNA was hybridized to Illumina Human HT-12 v4 Expression BeadChips. Genome Studio software (2010.v3) was used for obtaining fluorescent intensities for each probe. The probes were processed using nonparametric background correction, followed by quantile normalization with control and expression probes using the neqc function in the limma package (R v2.14.0) (22). The probes with detection *P* values <0.01 were considered expressed and used for subsequent analysis. Differential gene expression analyses to compare overexpression of miR-21-3p or miR-27a-5p with negative control were performed using Patterns of Gene Expression (PaGE v5.1.6) (23). Genes were considered to be differentially expressed if their expression changed at least 25% with <5% false discovery rate (FDR). The list of differentially expressed genes was interrogated for statistically significant over-represented biological themes using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (24). Gene expression data have been deposited in Gene Expression Omnibus with the accession number GSE48006.

### miRNA and mRNA RT-qPCR

miRNA reverse transcription was performed from total RNA samples that retained the small RNA fraction. The Applied Biosystems Taqman microRNA Reverse Transcription Kit was used to generate cDNA. miRNAs and small RNAs were quantified using the following Taqman assays: RNU44 (Assay ID 001095), hsa-miR-21-5p (Assay ID 000397), hsa-miR-21-3p (Assay ID 002438), hsa-miR-23a (Assay ID 000399), hsa-miR-24 (Assay ID 000402), hsa-miR-27a-3p (Assay ID 000408), and hsa-miR-27a-5p (Assay ID: 002445). mRNA reverse transcription was performed using the Taqman Reverse Transcription Kit. Quantitative PCR (qPCR) for miRNAs and mRNAs was performed using the Roche Probes 480 Master Mix or KAPA SYBRFast Master Mix, respectively, in a Roche LightCycler 480 instrument.

### Western blots

Cytosolic and nuclear fractions of total cellular protein were isolated using the Thermo Fisher Scientific NE-PER Nuclear and Cytoplasmic Extraction Kit. Isolates were run on 4–12% Bis-Tris gels and transferred onto polyvinylidene difluoride membranes. Antibodies used were p65 (ab7970, Abcam) at 1:250, Lamin A/C (sc-6215 and sc-7292, Santa Cruz) at 1:500, and  $\beta$ -Actin (5125, Cell Signaling) at 1:250 dilutions.

### 3' UTR assay

The 3' UTRs for all indicated genes were cloned into a Promega psiCHECK-2 plasmid. Fifteen nanograms per well 3' UTR plasmid

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

and 100 ng/well of  $\beta$ -galactosidase plasmid as filler were transfected along with miRNA mimic into HEK293 cells grown in 48-well plates for 24 h. Luciferase activity was measured using the Dual-Luciferase Reporter Assay System from Promega (E1960). The amount of Firefly luciferase activity was normalized to Renilla luciferase activity to account for transfection efficiency in each well.

#### Statistical analysis

Microarray differential expression was determined using a permutation-based method (23). In order to assess the statistical significance of changes in gene expression or luciferase activity, a two-sided Student's *t*-test with unequal variance was used.

## RESULTS

### miRNAs expressed in HAECs

We sequenced four small RNA libraries prepared using total RNA from HAECs from two separate donors treated with media alone or media containing 40  $\mu$ g/ml Ox-PAPC for 4 h. We obtained 15–26 million reads per library (supplementary Table 1). Between 40% and 62% of the reads mapped to genomic regions that overlapped with known mature miRNA sequences (mirBase v19) (25). To account for differences between sequencing depth of the libraries, we normalized the miRNA expression levels by calculating the percentage of reads that mapped to a specific miRNA with respect to the total number of reads that mapped to all known miRNAs for each sample. We detected between 386 and 583 miRNAs expressed per library, representing a total of 618 unique miRNAs expressed in HAECs (supplementary Table 1 and supplementary File). Of the 618 miRNAs identified, we found that 18 miRNAs were expressed at an abundance >1% of total miRNA reads. miRs-21-5p and -126-3p consisted of 40% of all the mapped sequencing reads. These findings indicate that while many miRNAs are expressed, only a small fraction constitutes the majority of the miRNA pool in HAECs (Fig. 1; supplementary File).

In order to validate the deep-sequencing results, we used miRNA-specific RT-qPCR to assay the expression levels of seven miRNAs whose abundance levels varied over a 10,000-fold range according to our deep-sequencing results. The

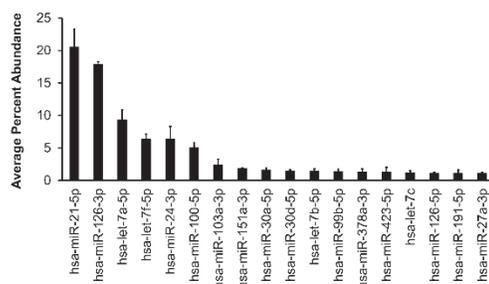


Fig. 1. Highly expressed HAEC miRNAs. miRNAs with average percent abundance >1% according to next-generation sequencing results are shown. Bars indicate average  $\pm$  SEM in four libraries.

correlation between the deep-sequencing and RT-qPCR results was highly significant (Pearson's  $R = 0.91$ ,  $P$  value =  $4.39 \times 10^{-6}$ ) suggesting that we were able to reliably quantify miRNA expression in HAECs using deep sequencing (supplementary Fig. 1).

### Effect of Ox-PAPC treatment on endothelial miRNAs

Oxidized glycerophospholipids (oxGPs) are known to contribute to numerous disease processes including atherosclerosis. Ox-PAPC is the major glycerophospholipid oxidized in minimally modified LDL that induces inflammatory and other responses in ECs. Treatment of cultured ECs with Ox-PAPC is known to change the expression of hundreds of genes in comparison to untreated ECs. Changes in miRNA expression levels as a response to environmental stimuli is known to modulate the expression of hundreds of genes; therefore, we hypothesized that Ox-PAPC may act to alter EC gene expression through changes in miRNA expression.

To assess the effects of Ox-PAPC treatment on EC miRNA expression, we calculated the fold change in abundance of miRNAs with respect to control cells. We observed that miRs-21-3p and -27a-5p were induced by 3.36- and 4.04-fold in response to Ox-PAPC treatment, respectively (Table 1). Mature miRNAs arise from Dicer processed pre-miRNA, a stem-loop precursor transcript that can produce two mature miRNAs, one from each arm (10). Although miRs-21-3p and -27a-5p are upregulated in response to Ox-PAPC, the expression of miRs-21-5p and -27a-3p, which arise from the same pre-miRNA transcript, was not altered by Ox-PAPC treatment (Table 1).

To determine the time course of miRNA induction, we measured the expression of miRs-21-3p and -27a-5p at 2, 4, and 8 h using miRNA-specific RT-qPCR. Consistent with the deep-sequencing results, we found that miRs-21-3p and -27a-5p were upregulated 2.0- and 2.5-fold, respectively at 4 h (Fig. 2A). In addition, we observed that miRs-21-3p and -27a-5p were upregulated 2- and 3-fold in response to Ox-PAPC as early as 2 h. By 8 h, miRs-21-3p and -27a-5p were still upregulated 2- and 3.5-fold, respectively.

Several distinct miRNAs can be expressed as a cluster from a single primary transcript. miR-27a-5p is transcribed from an miRNA gene cluster including the genes *MIR23A* and *MIR24-2*. To determine whether Ox-PAPC specifically induced the expression of miRs-21-3p and -27a-5p, and rather than all miRNAs that arise from the same primary transcripts, we measured the expression of additional miRNAs at the three time points. Among the six miRNAs measured, only miRs-21-3p and -27a-5p showed induction upon treatment (Fig. 2A) suggesting that the regulation of miRs-21-3p and -27a-5p expression by Ox-PAPC occurs after these miRNAs are processed to mature form.

To determine whether induction of miRs-21-3p and -27a-5p in ECs was specific for Ox-PAPC, we treated HAECs with alternative inflammatory stimuli: IL-1 $\beta$  and TNFC. Exposure to TNFC is known to induce the expression of miRs-155-5p and -31-5p in human umbilical vein endothelial cells (HUVECs) (26). Therefore, to confirm TNFC activity, we measured the expression of these two miRNAs to serve as

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

TABLE 1. Abundance of miR-21 and miR-27a (percentage) in HAECs

	Donor 1		Donor 2		Fold Change
	Control	Ox-PAPC	Control	Ox-PAPC	Ox-PAPC versus Control
miR-21-5p	$1.64 \times 10^1$	$1.56 \times 10^1$	$2.35 \times 10^1$	$2.68 \times 10^1$	1.04
miR-21-3p	$5.15 \times 10^{-3}$	$1.80 \times 10^{-2}$	$1.32 \times 10^{-2}$	$4.28 \times 10^{-2}$	3.37
miR-27a-3p	$1.11 \times 10^0$	$9.31 \times 10^{-1}$	$1.21 \times 10^0$	$1.34 \times 10^0$	0.97
miR-27a-5p	$2.43 \times 10^{-2}$	$1.06 \times 10^{-1}$	$2.36 \times 10^{-2}$	$8.77 \times 10^{-2}$	4.03

positive controls. We observed that miR-21-3p was induced *rv*4- and 2-fold by IL-1 $\beta$  and TNFC, respectively, while miR-27a-5p was induced *rv*2.5- and 3.5-fold by IL-1 $\beta$  and TNFC, respectively (Fig. 2B). These results suggest that miRs-21-3p and -27a-5p may be induced as a general response to inflammatory stimuli in HAECs.

#### Identification of biological pathways regulated miRs-21-3p and -27a-5p in ECs

A small number of studies have characterized mRNA targets of miRs-21-3p and -27a-5p in immune cells such as eosinophils and natural killer (NK) cells and cancers such as hepatocellular carcinoma and head and neck squamous cell carcinoma (27–30). However, how miRs-21-3p and -27a-5p regulate gene expression in ECs is not known. Therefore, we sought to identify the targets and biological pathways modulated by these two miRNAs in ECs using an unbiased approach. Based on previous dose response experiments with miR-21-3p and miR-27a-5p miRNA mimics, we transfected HAECs from two additional donors with 1 nM mimic leading to 5- to 30-fold induction of miR-21-5p and 15- to 30-fold induction of miR-27a-5p, respectively (supplementary Figs. II and III). Twenty-four hours after transfection, we isolated the total RNA and performed whole genome transcript profiling. Because miRNAs often function to degrade mRNAs, we focused on the transcripts that were downregulated in response to the transient overexpression.

We found that miRs-21-3p and -27a-5p decreased the expression of 1,037 and 1,138 genes >25% at 5% FDR, respectively. Of the 1,253 genes that were downregulated

by either of the miRNAs, 922 of them were common, suggesting the regulation of similar biological pathways by these two miRNAs in ECs. While a single miRNA can target several hundred mRNAs, it is likely that most of the downregulation of the genes in HAECs is due to secondary effects. miRNAs typically bind to the 3' UTRs of their mRNA targets leading to transcript degradation (10). We examined all 1,253 genes downregulated by either miRNA using the miRNA target prediction algorithm miRANDA (31). We found that 471 and 411 genes contained predicted binding sites for miRs-21-3p or -27a-5p, respectively. Of those genes with predicted miRNA binding sites, 206 genes were predicted targets for both miRNAs.

Using DAVID, we searched for Gene Ontology enrichment in the genes downregulated by miRNA overexpression. We observed significant enrichment for genes involved in the defense response, inflammatory response, and response to wounding functional categories (Table 2). Within the defense response category, genes belonging to the NF- $\kappa$ B signaling pathway were highly represented comprising *rv*30% and 26% percent of genes downregulated by miRs-21-3p and -27a-5p, respectively. Therefore, we hypothesized that modulation of miRs-21-3p and -27a-5p expression in ECs would lead to functional consequences for NF- $\kappa$ B signaling.

#### miRs-21-3p and -27a-5p regulate p65 nuclear translocation

The p65, one of the NF- $\kappa$ B subunits, is sequestered in the cytoplasm by the I $\kappa$ B proteins in an inactive state (32). Once the signaling cascade is activated, the I $\kappa$ B proteins are phosphorylated by the IKK family of kinases, and p65 is free

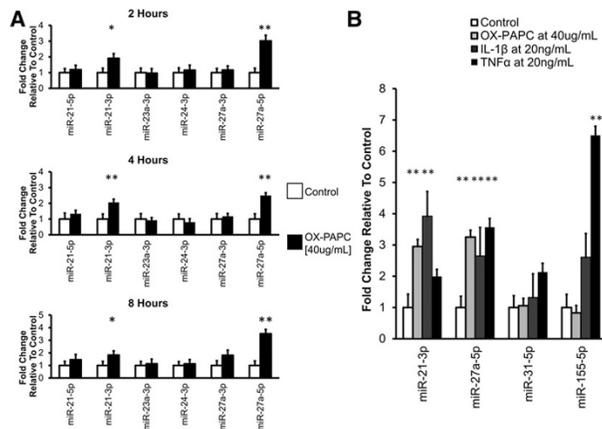


Fig. 2. Expression of select miRNAs in HAECs in response to various stimuli. A: miRNA expression in HAECs over time in response to Ox-PAPC treatment. HAECs were incubated with either control media or media containing Ox-PAPC for the indicated durations. miRNA levels were quantified using RT-qPCR. Plots show average  $\pm$  SEM. N = 8–9. B: miRNA expression in HAECs treated with inflammatory stimuli. HAECs were treated with either control media or Ox-PAPC for 4 h, or IL-1 $\beta$  or TNFC for 2 h. miRNA levels were quantified using RT-qPCR. Plots show average  $\pm$  SEM. N = 6. \*\*  $p < 0.01$ .

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

TABLE 2. Functional enrichment of Gene Ontology categories of genes downregulated by miR-21-3p and miR-27a-5p

Gene Ontology Category	miR-21-3p Overexpression			miR-27a-5p Overexpression		
	Number of Genes	Enrichment <i>P</i>	Bonferroni - Corrected Enrichment <i>P</i>	Number of Genes	Enrichment <i>P</i>	Bonferroni - Corrected Enrichment <i>P</i>
Defense response	119	$6.04 \times 10^{-15}$	$2.58 \times 10^{-11}$	130	$4.99 \times 10^{-15}$	$2.32 \times 10^{-11}$
Inflammatory response	74	$5.77 \times 10^{-12}$	$2.49 \times 10^{-8}$	80	$7.43 \times 10^{-12}$	$3.45 \times 10^{-8}$
Response to wounding	107	$1.21 \times 10^{-10}$	$5.20 \times 10^{-7}$	119	$3.01 \times 10^{-11}$	$1.40 \times 10^{-7}$
Response to lipopolysaccharide	27	$8.33 \times 10^{-6}$	$3.01 \times 10^{-1}$	28	$3.28 \times 10^{-4}$	$7.82 \times 10^{-1}$
Regulation of cell death	173	$3.65 \times 10^{-5}$	$1.46 \times 10^{-1}$	193	$5.41 \times 10^{-5}$	$2.22 \times 10^{-1}$
Regulation of I $\kappa$ B kinase/NF- $\kappa$ B cascade	36	$4.53 \times 10^{-4}$	$8.58 \times 10^{-1}$	34	$1.94 \times 10^{-2}$	$1.00 \times 10^0$

to translocate to the nucleus. Based on the significant decrease in NF- $\kappa$ B signaling related genes in response to miRNA overexpression, we hypothesized that overexpression of miRs-21-3p and -27a-5p would impair p65 nuclear translocation. Previous work in HeLa cells using a luciferase reporter construct containing three tandem NF- $\kappa$ B binding sites has strongly suggested that Ox-PAPC does not activate NF- $\kappa$ B signaling (33). We observed that treatment of HAECs with Ox-PAPC did not induce p65 nuclear translocation but appeared to decrease basal nuclear translocation (supplementary Fig. IV). Consequently, to determine the effect of miRs-21-3p and -27a-5p on p65 nuclear translocation, we chose to treat HAECs overexpressing miRs-21-3p and -27a-5p with IL-1 $\beta$  a known activator of NF- $\kappa$ B signaling.

As expected, p65 translocated to the nucleus in response to IL-1 $\beta$  treatment in cells transfected with control miRNA (Fig. 3A). We observed a significant decrease in p65 nuclear translocation following overexpression of miRs-21-3p and -27a-5p (Fig. 3A, B). We observed a lesser decrease in p65 nuclear translocation with IL-1 $\beta$  treatment of HAECs following knockdown of miRs-21-3p and -27a-5p expression with miRNA inhibitors (Fig. 3C, D).

Altered p65 nuclear translocation leads to changes in the expression of downstream targets of NF- $\kappa$ B signaling. We measured the expression of E-selectin and vascular cell adhesion molecule-1 (VCAM-1), direct targets of NF- $\kappa$ B, in HAECs overexpressing the two miRNAs. In control cells, E-selectin and VCAM-1 expression were induced >70-fold in response to IL-1 $\beta$  treatment. As expected, we observed impaired induction of E-selectin and VCAM-1 following miRNA overexpression (Fig. 3E, F).

As knockdown of the two miRNAs appeared to also alter p65 nuclear translocation, we measured the expression of E-selectin and VCAM-1 in ECs transfected with miRNA inhibitors. We found that in comparison to control cells, knockdown of miR-27a-5p had no significant change on the expression of E-selectin or VCAM-1 (supplementary Fig. V). In comparison, knockdown of miR-21-3p led to an increased expression of both E-selectin and VCAM-1, despite impaired p65 nuclear translocation (supplementary Fig. V).

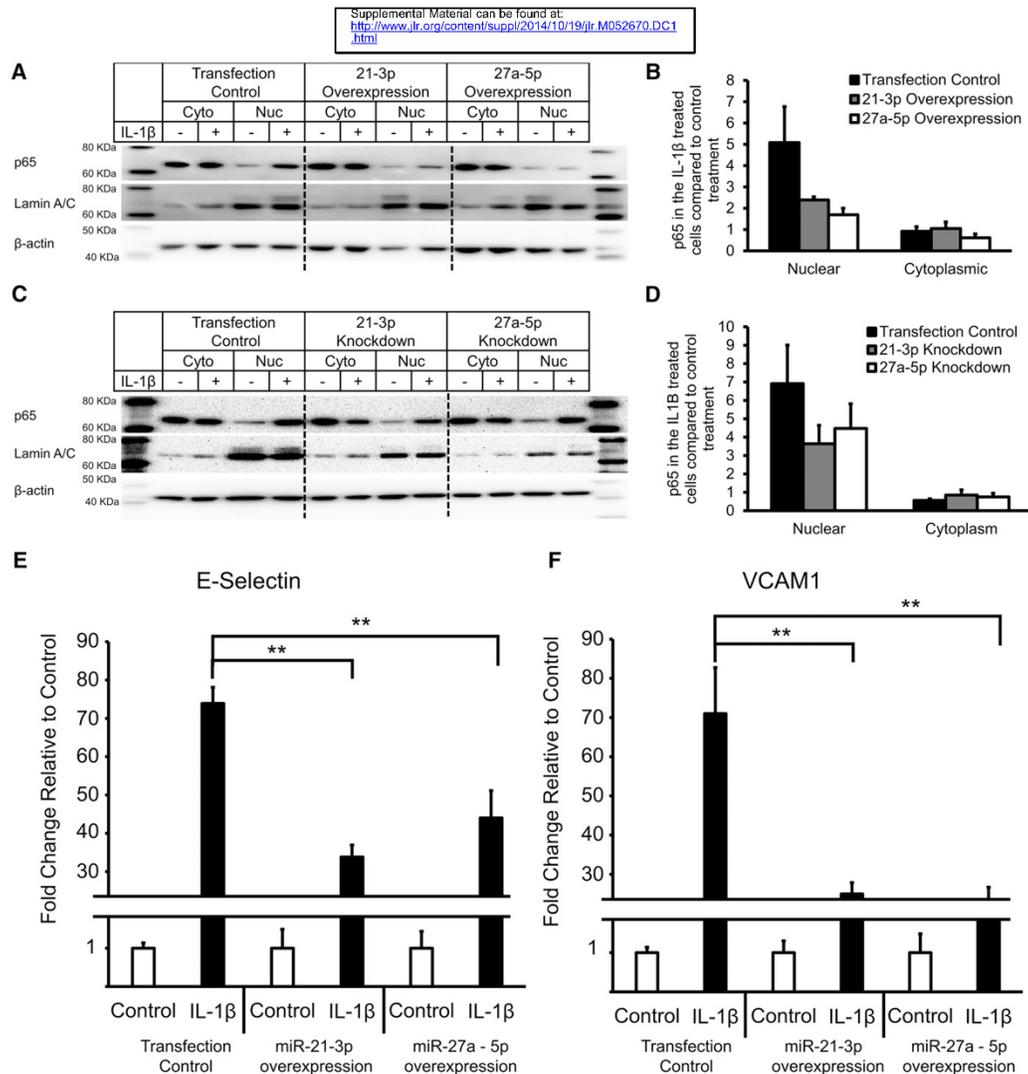
#### Regulation of NF- $\kappa$ B signaling through 3' UTR targeting of NF- $\kappa$ B genes

We used two distinct approaches to identify candidate target genes for miRs-21-3p and -27a-5p (supplementary Fig. VI). In our first approach, we focused on the differentially

downregulated genes in our microarrays studies. There were 1,037 and 1,138 downregulated genes in response to miRs-21-3p and -27a-5p overexpression, respectively. 119 and 130 of these genes belonged to the defense response Gene Ontology functional category, respectively. The two gene lists along with known NF- $\kappa$ B pathway genes that were found to be downregulated by miRs-21-3p or -27a-5p overexpression were then analyzed for putative binding sites for either miRs-21-3p or -27a-5p as predicted by the miRANDA algorithm (supplementary Fig. VI). This approach generated 26 preliminary candidate genes that were then assayed using RT-qPCR in an additional HAEC donor overexpressing miRs-21-3p or -27a-5p. We were able to confirm downregulation for 15 of the 26 genes that contained putative binding sites for either of the two miRNAs (*P* < 0.1) (supplementary Fig. VII).

In the second approach, we queried the whole genome for predicted targets of miRs-21-3p and -27a-5p using miRWALK, a database of miRNA prediction algorithms (36). Genes that showed predicted target sites for either miRs-21-3p or -27a-5p by a minimum of three algorithms were then searched in the literature for known associations with NF- $\kappa$ B signaling or known relevance to EC biology. This resulted in 12 putative targets for miR-21-3p and 8 putative targets for miR-27a-5p. Five genes out of the 20 genes were selected at random to be validated as miRNA targets using the 3' UTR assay. The combination of these two approaches resulted in 20 candidate target genes representing three distinct miRs-21-3p and -27a-5p predicted target types: joint miRs-21-3p and -27a-5p target, miR-21-3p target only, and miR-27a-5p target only (Table 3; supplementary Fig. VI).

Upon close inspection of these 20 candidate target genes, we found they could be classified into three major subgroups related to their location/role in the NF- $\kappa$ B signaling cascade. Eight genes (*IL1R1*, *TICAM2*, *TLR3*, *UBE2N*, *BCL10*, *MAP2K7*, *TNFAIP3*, and *TNIP1*) have been shown to act upstream of p65 nuclear translocation. The remaining 12 genes involved processes downstream of p65 nuclear translocation and could be broken up into two additional groups. Three genes (*RELA*, *RFC1*, and *CEBPB*) were found to act at the site of NF- $\kappa$ B transcriptional binding sites. The remaining nine genes (*GCH1*, *CBX4*, *HMOX1*, *BCL2*, *CX3CL1*, *IL12A*, *HMGB1*, *ICOSLG*, and *TAP1*) were known downstream targets of NF- $\kappa$ B. These nine genes based on known biological function could be classified into two main subcategories: cell survival and proliferation



**Fig. 3.** miR-21-3p and miR-27a-5p inhibit NF- $\kappa$ B signaling in HAECs. **A:** HAECs were transfected with 1 nM control or miR-21-3p or miR-27a-5p mimic for 24 h and then treated with 20 ng/ml IL-1 $\beta$  for 2 h. Representative Western blot of both cytosolic and nuclear protein fractions from HAECs. **B:** Quantification of p65 nuclear translocation in three donors. Bars show average  $\pm$  SD.  $N = 3$ . **C:** HAECs were transfected with 50 nM control or miR-21-3p or miR-27a-5p inhibitor for 24 h and then treated with 20 ng/ml IL-1 $\beta$  for 2 h. Representative Western blot of both cytosolic and nuclear protein fractions from HAECs. **D:** Quantification of p65 nuclear translocation in three donors. Bars show average  $\pm$  SD.  $N = 3$ . **E and F:** After 24 h transfection with control and mimics, cells were treated with 2 ng/ml of IL-1 $\beta$  for 2 h; gene expression was quantified using RT-qPCR. Bars represent the average  $\pm$  SEM.  $N = 6$ .  $**P < 0.01$ .

(*GCH1*, *CBX4*, *HMOX1*, and *BCL2*) and immune response/inflammation (*CX3CL1*, *IL12A*, *HMGB1*, *ICOSLG*, and *TAP1*). Thus, it is likely that miRs-21-3p and -27a-5p act at many points in the signaling cascade to regulate the extent of NF- $\kappa$ B signaling.

miRNAs regulate gene expression through direct interactions with the 3' UTR of the target mRNA. Therefore, to determine direct miRNA/mRNA interactions by

miRs-21-3p and -27a-5p with their putative target genes, we utilized 3' UTR luciferase reporter assay. The 3' UTRs of the 20 genes were cloned to create fusion transcripts with the firefly luciferase in the psiCHECK-2 plasmid. The plasmids containing the candidate gene 3' UTRs were then cotransfected in HEK293 cells with 5 nM mimics of miR-21-3p, miR-27a-5p, or cel-miR-67 (a negative control) to overexpress the miRNAs of interest.

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

TABLE 3. Putative gene targets of miR-21-3p and miR-27a-5p for validation

Gene Symbol	Gene Name	miR-21-3p Binding Site?	miR-27a-5p Binding Site?
<i>BCL2</i>	B-cell CLL/lymphoma 2	Ye	Ye
<i>HMOX1</i>	Heme oxygenase (decycling) 1	s	s
		Ye	Ye
		s	s
<i>ICOSLG</i>	Inducible T-cell costimulator ligand	Yes	Yes
<i>IL1R1</i>	Interleukin 1 receptor, type 1	Yes	Yes
<i>TICAM2</i>	Toll-like receptor adaptor molecule 2	Yes	Yes
<i>UBE2N</i>	Ubiquitin-conjugating enzyme E2N	Yes	Yes
<i>BCL10</i>	B-cell CLL/lymphoma 10	Yes	No
<i>CBX4</i>	Chromobox homolog 4	Yes	No
<i>CEBPB</i>	CCAAT/enhancer binding protein (C/EBP) beta	Yes	No
<i>CX3CL1</i>	Chemokine (C-X3-C motif) ligand 1	Yes	No
<i>GCH1</i>	GTP cyclohydrolase 1	Yes	No
<i>HMGB1</i>	High mobility group box 1	Yes	No
<i>IL12A</i>	Interleukin 12A	Yes	No
<i>TAP1</i>	Transporter 1, ATP-binding cassette, subfamily B	Yes	No
<i>TLR3</i>	Toll-like receptor 3	Yes	No
<i>TNFAIP3</i>	Tumor necrosis factor, alpha-induced protein 3	Yes	No
<i>RELA</i>	V-rel reticuloendotheliosis viral oncogene homolog A/p65	No	Yes
<i>RFC1</i>	Replication factor C (activator 1) 1, 145 kDa	No	Yes
<i>TNIP1</i>	TNFAIP3 interacting protein 1	No	Yes

Of the 20 candidate genes, 16 genes contained predicted miR-21-3p sites in their 3' UTR. For 11 of 16 genes (*BCL2*, *ICOSLG*, *TICAM2*, *UBE2N*, *BCL10*, *CBX4*, *CEBPB*, *HMGB1*, *IL12A*, *TAP1*, and *TLR3*), we observed significant down-regulation of luciferase activity following cotransfection

with miR-21-3p mimic (Fig. 4A). Seven 3' UTR plasmid constructs representing the three miR-21-3p candidate target genes (*GCH1*, *TNFAIP3*, and *CX3CL1*) and three joint miR-21-3p and miR-27a-5p candidate target genes (*HMOX1*, *MAP2K7*, and *IL1R1*) did not show decreased

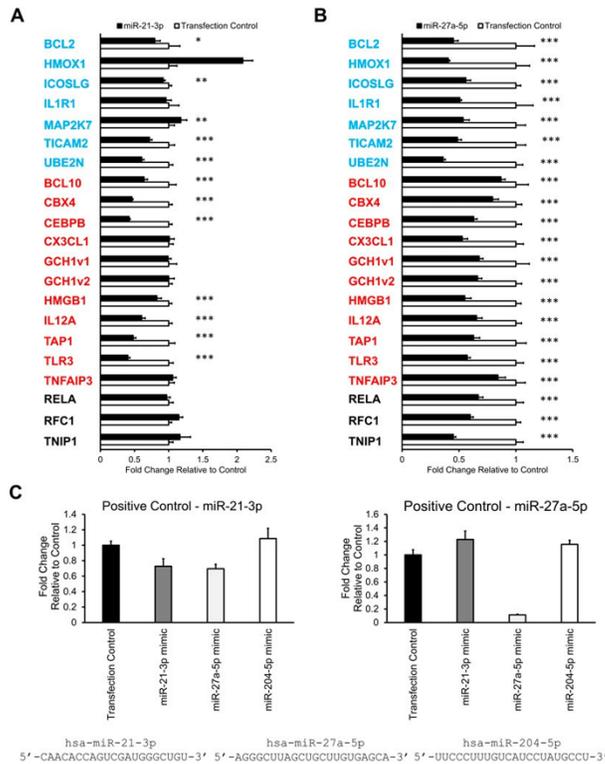


Fig. 4. Identification of candidate target genes of miR-21-3p and miR-27a-5p. A and B: For each gene, a plasmid containing the 3' UTR was cotransfected with miRNA mimic at 5 nM concentration into HEK293 cells. Results for the miR-21-3p mimic transfected cells (A) and for the miR-27a-5p mimic transfected cells (B). Bars represent average  $\pm$  SD. N = 6. A and B: Blue color indicates genes that are predicted to be miR-21-3p and miR-27a-5p targets. Red color indicates genes that are predicted to be miR-21-3p targets. Black color indicates genes that are predicted to be miR-27a-5p targets. GCH1v1 and GCH1v2 represent two different 3' UTR isoforms of the same gene *GCH1*. C and D: 3' UTR luciferase assay positive controls for miR-21-3p and miR-27a-5p. For each miRNA, a positive control plasmid containing the full-length complementary sequence of the miRNA in triplicate was generated. These positive control plasmids were cotransfected with either mimic control (transfection control), miR-21-3p, miR-27a-5p, or miR-204-5p (an unrelated miRNA) into HEK293 cells. C: Results from the miR-21-3p positive control and from the miR-27a-5p positive control. Bars represent average  $\pm$  SD. N = 6. Asterisks indicate significant downregulation following miRNA overexpression. \*  $P < 0.1$ , \*\*  $P < 0.05$ , \*\*\*  $P < 0.01$ .

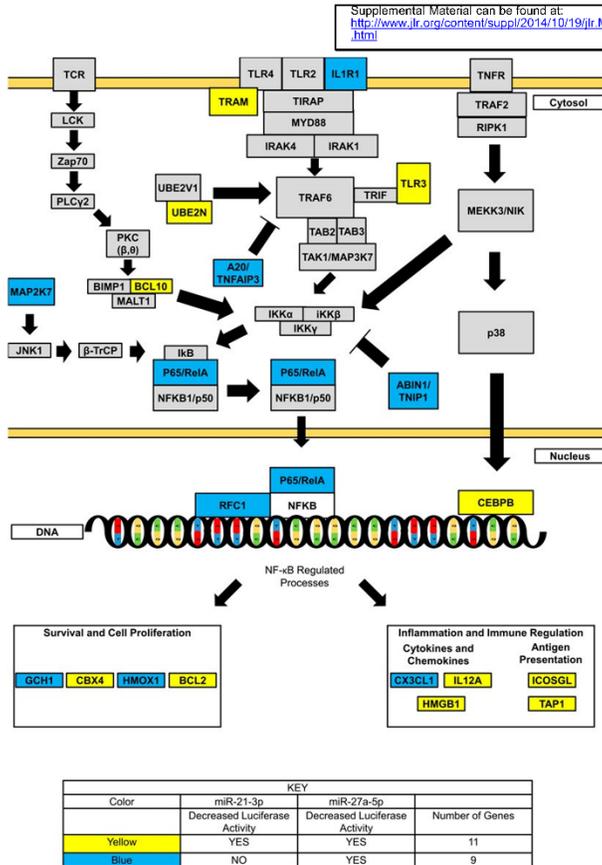


Fig. 5. miRs-21-3p and -27a-5p affect the expression of multiple genes in the NF-κB signaling cascade. The 20 candidate genes tested in the 3' UTR luciferase assay were overlaid onto the NF-κB signaling cascade. Yellow color indicates genes that were shown to be regulated by both miRs-21-3p and -27a-5p overexpression using 3' UTR luciferase assay. Blue color indicates genes that were shown to be regulated by miR-27a-5p overexpression, but not miR-21-3p overexpression using 3' UTR luciferase assay.

luciferase activity following miR-21-3p overexpression. Numerous factors can contribute to the strength of miRNA-3' UTR interactions including location of predicted miRNA binding site in the 3' UTR, extent of complementarity between miRNA seed sequence and target 3' UTR, and total extent complementarity between the full-length miRNA sequence and target 3' UTR (10). We hypothesized that variation in one of these three factors was shared by the six candidate genes leading to a lack of response following miR-21-3p overexpression. However, upon closer examination, we were unable to find significant similarity among the three factors in the six candidate genes.

In contrast, 10 out of 20 candidate genes contained predicted miR-27a-5p interaction sites as determined by miRANDA. Using the luciferase reporter assay, we found that all 10 predicted candidate genes showed decreased luciferase activity in response to miR-27a-5p overexpression (Fig. 4B). Furthermore, we observed that all 10 candidate genes predicted by miRANDA to contain only miR-21-3p binding sites also showed downregulation in response to miR-27a-5p, independent of their response to miR-21-3p overexpression (Fig. 4A, B).

miRNA recognition of its target 3' UTRs is primarily mediated by complementary in the target 3' UTR to miRNA

seed sequence, positions 2–8 of the mature miRNA. However, studies using Argonaute (Ago)-bound miRNA/mRNA pairings suggest that up to 45% miRNA targets lack a perfect seed sequence match (37). Therefore, we tested the hypothesis that there was a nonspecific interaction between miR-27a-5p and miR-21-3p target 3' UTR sites. We generated positive control luciferase plasmids for both miR-21-3p and miR-27a-5p by inserting sequences that were complementary to the full sequence of miRs-21-3p or -27a-5p into the luciferase plasmid. We then cotransfected the positive control miRNA target plasmids with the miRNA mimics and then measured luciferase activity 48 h after transfection. As a negative control we cotransfected the positive control miRNA target plasmids with miR-204-5p mimic, an unrelated miRNA that shows no sequence similarity to either miR-21-3p or miR-27a-5p (Fig. 4C). As expected, both the miRs-21-3p and -27a-5p positive controls showed no downregulation of luciferase activity in response to miR-204-5p overexpression in addition to downregulation of luciferase activity in response to overexpression of their respective miRNA (Fig. 4C).

There was no effect of miR-21-3p or miR-204-5p overexpression on the luciferase construct harboring the miR-27a-5p complementary sequence (Fig. 4C). However,

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

we observed a decrease in luciferase activity of the miR-21-3p positive control plasmid with miR-27a-5p overexpression and not with miR-204-5p overexpression (Fig. 4C). This finding suggested that there is possible direct interaction between miR-21-3p target sites and miR-27a-5p that is not predicted by in silico prediction methods (supplementary Table II). Therefore, to confirm that miR-27a-5p can indeed interact with predicted miR-21-3p sites in 3' UTRs of genes, we mutated the predicted miR-21-3p site and measured luciferase activity in response to miR-21-3p and miR-27a-5p overexpression. At random we selected the 3' UTR of the *CBEPB* gene and mutated the miR-21-3p predicted target site (supplementary Fig. VIII B). Upon overexpression of the two miRNAs, we found that luciferase activity decreased ~60% and 40% in response to miRs-21-3p and -27a-5p overexpression, respectively (supplementary Fig. VIII A). In comparison, mutation of the predicted miR-21-3p in the *CBEPB* 3' UTR abolished the effects of miRs-21-3p and -27a-5p overexpression on luciferase activity. These findings strongly suggest that miR-27a-5p can interact with predicted miR-21-3p sites in 3' UTR leading to downregulation of mRNA expression (supplementary Fig. VIII).

## DISCUSSION

The relevance of the effects of oxGPs on endothelial function in various pathologies, such as acute inflammation, lung injury, and atherosclerosis, has been extensively documented (38). Here, we identified a role for two miRNAs, miRs-21-3p and -27a-5p, in regulating the response of HAECs to oxGPs. Using an unbiased global approach, we determined that the two miRNAs regulate the NF- $\kappa$ B signaling pathway in ECs in response to additional inflammatory stimuli including IL-1 $\beta$ . Using the 3' UTR luciferase assay system, we show interactions between the two miRNAs with multiple transcripts in the NF- $\kappa$ B signaling cascade suggesting that these miRNAs act as fine-tuning regulators of NF- $\kappa$ B signaling.

### Profile of miRNA-expressed in ECs

The endothelium contributes to the regulation of numerous physiological functions including the vasomotor tone, hemostatic balance, and inflammation. Here we present a catalog of miRNAs expressed in HAECs as identified by next-generation sequencing of the small RNAs. We identified a total of 618 unique miRNAs expressed in HAECs. Previous studies characterizing miRNA expression in ECs from seven different vascular beds expression using miRNA microarrays identified 166 EC miRNAs (39). Our study identified 474 HAEC miRNAs in addition to replicating 144 of the 166 previously reported EC miRNAs. The greater number of miRNAs detected in our study compared with the previous study can in part be attributed to the use of deep sequencing for detection. In contrast to microarray methods, deep sequencing contains a greater dynamic range for quantification of miRNA expression as well as increased accuracy in distinguishing between similar miRNAs and allowing for identification of novel miRNAs (40).

Previously reported deep-sequencing results identified 427 known and novel miRNAs expressed in HUVECs (41). Our study identified 369 of these 427 miRNAs. While we cannot rule out the differences in sequencing depth and data analysis between the two studies, these results suggest that although the majority of miRNAs are common between HAECs and HUVECs, differences in miRNA expression may exist based on the vascular bed.

Characterization of miRNA-regulated biological processes in ECs through identification of miRNA/mRNA interactions has shown that miRNAs can act as key regulators of numerous physiological processes in ECs including angiogenesis, lineage commitment, and inflammation (11, 42, 43). Of the 18 highly expressed EC miRNAs identified in our study, 11 have been previously studied in ECs. Seven of the 11 previously expressed miRNAs have been implicated as regulators of angiogenesis. miRs-126-5p, -126-3p, -27a-3p, -103a-3p, -30a-5p, and -let-7f are proangiogenic and target known inhibitors of angiogenesis, including the genes *SPRED1*, *DLK1*, *SEMA6*, *TSP-1*, and *DLL4* to promote vessel development (11, 44-47). miR-24-3p, through regulation of the transcription factor of *GATA2* and *PAK4*, acts to inhibit angiogenesis (48).

### EC miRNAs in the immune response

In addition to angiogenesis, highly expressed EC miRNAs have also been shown to act as regulators of the inflammatory process. miR-126-3p, in addition to its role as a modulator of angiogenesis, is a known anti-inflammatory miRNA through direct repression of the leukocyte adhesion molecule VCAM-1 (49). In addition, many miRNAs that are expressed at moderate to low levels in our data set have also been shown to regulate inflammation in ECs. The moderately expressed miRs-17-5p and -31-5p directly inhibit the transcription of leukocyte adhesion molecules intracellular adhesion molecule-1 and E-selectin, respectively (26). Along with miR-126-3p, this provides evidence for miRNAs as direct suppressors of inflammation in ECs (26, 49). In addition to direct suppression of the expression of inflammatory molecules, anti-inflammatory EC miRNAs also act to regulate major signaling pathways. For example, miRs-181b-5p, -155-5p, and -10a-5p act to suppress inflammation through inhibition of NF- $\kappa$ B signaling (15, 43). Our studies identified two EC miRNAs that act as inhibitors of inflammation in ECs, miRs-21-3p and -27a-5p, through regulation of NF- $\kappa$ B signaling.

### Regulation of NF- $\kappa$ B signaling by miRs-21-3p and -27a-5p in ECs

NF- $\kappa$ B signaling is one of the key pathways in the activation of the immune response in ECs. Various environmental cues, including cytokines, bacterial infections, and hemodynamic forces, can activate NF- $\kappa$ B signaling in ECs leading to the increased expression of leukocyte adhesion molecules, cytokines, and chemokines (50-52). In ECs, miRNAs have been shown to suppress the activation of NF- $\kappa$ B at various points in the signaling cascade. A key site of miRNA-mediated suppression of NF- $\kappa$ B signaling is through degradation of the kinases responsible for phosphorylating

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

IKB repressor proteins. miRs-10a-5p, -155-5p, -199a-5p, -223-3p, -15a-5p, and -16-5p all function in this manner to repress NF- $\kappa$ B signaling (15, 53–55). In addition, EC miRNAs can also repress NF- $\kappa$ B signaling by inhibiting the translocation of the NF- $\kappa$ B subunits to the nucleus, such as miR-181b-5p-mediated transcriptional repression of importin-C3, an adaptor protein functioning in nuclear protein import (43).

Based on our microarray data, we predicted that Ox-PAPC induced an increase in the levels of miRs-21-3p and -27a-5p to suppress NF- $\kappa$ B signaling in ECs. Previous studies from several groups have demonstrated that in HAECs and HUVECs, Ox-PAPC minimally activates or does not activate the expression of many NF- $\kappa$ B targets strongly induced by lipopolysaccharide (LPS), TNFC, or IL-1 $\beta$  such as E-selectin and VCAM-1 (34, 35). Using a luciferase construct containing three copies of the NF- $\kappa$ B responsive element, it has previously shown that unlike TNFC, Ox-PAPC treatment does not induce activation of NF- $\kappa$ B (33). Rather, these and other studies indicate that Ox-PAPC is a strong inhibitor of activation of NF- $\kappa$ B targets by LPS in all EC types (35). Our studies suggest that the induction of miRs-21-3p and -27a-5p by Ox-PAPC can be a possible mechanism that suppresses the activation of NF- $\kappa$ B signaling in ECs.

In contrast to Ox-PAPC, IL-1 $\beta$  and TNFC have also been shown to strongly induce synthesis of proinflammatory molecules by the NF- $\kappa$ B pathway but have well established feedback mechanisms to block long-term activation, including miRNAs. Our studies demonstrate that miRs-21-3p and -27a-5p are used to regulate NF- $\kappa$ B activation by these divergent activators of inflammation.

#### Identifying miRNA targets

In an attempt to identify direct targets of miRs-21-3p and -27a-5p, we used two in silico approaches. In our first approach, we identified candidate genes by focusing on genes whose expression was downregulated following miRNA overexpression. While this approach has been well established for the identification of miRNA targets, the large number of genes whose expression was altered by either miRs-21-3p or -27a-5p overexpression in our data set, even at a highly conservative statistical cutoff, necessitated candidate prioritization. In our second method, we utilized an in silico-based approach for identifying miRNA target genes by using multiple miRNA prediction programs to query the whole genome for predicted targets of miRs-21-3p and -27a-5p. In comparison to direct measurements of miRNA/mRNA interactions, the utilization of miRNA prediction algorithms such as miRANDA, which rely on base complementarity of the mature miRNA sequence in the 3' UTR, biased our final candidate gene list to genes that contained sequences in the 3' UTR complementary to miRs-21-3p and -27a-5p seed sequence (56).

While numerous methods have been developed for the identification of direct miRNA-mRNA interactions, the application of these methods is highly dependent on the miRNA-mRNA interaction being investigated. cross-linking immunoprecipitation-based methods, which utilize

immunoprecipitation of Ago proteins in the RNA-induced silencing complex complex in combination with cross-linking of bound direct miRNA-mRNA interactions, have the advantage of direct determination of in vivo miRNA-mRNA interactions. However, it should be noted that low-abundance target miRNA-mRNA interactions may not be captured regardless of their affinity of interaction with Ago (57). To validate our list of 20 candidate genes, we chose to use the well-established 3' UTR luciferase assay. One of the drawbacks of this method is its reliance on excessively high overexpression of both the target UTR and miRNA to determine affinity between the miRNA and its target.

Upon testing of these 20 candidate genes using the 3' UTR luciferase assay, we made two distinct observations. First, in testing miR-21-3p target sites, we were only able to confirm downregulation for 11 of the 16 predicted target genes, representing an rv69% concordance of miRNA/mRNA interaction predictions for miR-21-3p by miRANDA with experimental observation. In contrast, previous studies on the accuracy of miRNA prediction programs for predicting true miRNA/mRNA interactions have reported 20–49% concordance between experimental validation and in silico prediction for miRANDA (58, 59).

Second, we observed downregulation of all 10 putative miR-21-3p targets containing no predicted miR-27a-5p targets with miR-27a-5p overexpression. We further confirm this observation using a positive control luciferase plasmid containing only the complementary sequence of miR-21-3p. In addition, we do not observe this effect with miR-21-3p overexpression and miR-27a-5p target sites. Upon visual inspection of the mature sequences of miRs-21-3p and -27a-5p, we have identified two regions of limited sequence similarity that may contribute to the interaction of miR-27a-5p to miR-21-3p target sites. In mammals, canonical miRNA/mRNA interactions involve complementarity between the 3' UTR and the given miRNA's 5' end. The first region of limited region of sequence similarity involves bases 13 and 15–19 of miR-21-3p being identical to bases 1–5 of miR-27a-5p, which represents a large fraction of the miR-27a-5p seed sequence suggesting a possible mechanism for downregulation of miR-21-3p targets by miR-27a-5p.

Complementarity in other regions beyond the 2–8 base pair seed sequence at the 5' UTR has also been reported to mediate miRNA/mRNA interactions (10). One of these alternate miRNA/mRNA interactions, centered miRNA binding, involves the presence of 11 base perfect complementarity starting at base 3, 4, or 5 of the miRNA sequence to 3' UTR (60). In addition to similarity to miR-27a-5p seed sequence, between bases 8 and 17 of both miRNAs, bases 8, 9, 12, 14, and 16 are identical between the two miRNAs, suggesting an alternate interaction site between miR-21-3p sites and miR-27a-5p.

We show further evidence that miR-27a-5p can interact with predicted miR-21-3p sites as mutation of the predicted miR-21-3p site in the 3' UTR of *CBEPB* abolished the effects of both miRs-21-3p and -27a-5p overexpression on luciferase activity. These observations strongly suggest that these two miRNAs act to regulate gene expression in



Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

the NF- $\kappa$ B signaling pathway through both through down-regulation miRNA-specific target genes and, in a synergistic manner, through direct targeting of the same genes.

#### miRs-21-3p and -27a-5p modulate the extent of NF- $\kappa$ B signaling in ECs

miRNAs are biological rheostats that act to fine-tune gene expression to provide robustness and modulate the extent of activation of the signaling cascade (61). One mechanism of modulation of a signaling cascade is through direct repression of both activators and repressors, as seen in zebrafish miR-430 targeting both transforming growth factor  $\beta$  nodal agonist *squinty* and nodal antagonist *lefty* (62). In total, we found evidence for direct interactions between miRs-21-3p and/or -27a-5p for 20 candidate genes (Figs. 4, 5). Upon closer inspection of these 20 candidate genes, we found that several of these genes are either known repressors of NF- $\kappa$ B signaling, such as *TNFAIP3/A20* and *TNIP1/ABIN-1*, or activators of NF- $\kappa$ B signaling, such as *MAP2K7* (63). In addition, we found that 9 of the 20 candidate genes lie downstream of NF- $\kappa$ B signaling (Fig. 5). These two findings, taken in consideration with the experimental observations that both overexpression and knockdown of miRs-21-3p and -27a-5p impairs p65 nuclear translocation following IL-1 $\beta$  treatment, suggest a dual role for miRs-21-3p and -27a-5p interacting with both the activators and regulators of NF- $\kappa$ B signaling.

Activation of gene expression in response to IL-1 $\beta$  is complex and involves the activation of numerous signaling cascades including c-Jun N-terminal kinase, MAPK, and NF- $\kappa$ B (64). Following inhibition of miRs-21-3p and -27a-5p with miRNA inhibitors, we did not observe the expected decrease in expression of E-selectin and VCAM-1 given that loss of the miRNAs led to a decrease in p65 nuclear translocation. Treatment of HAECs with IL-1 $\beta$  following knockdown of miR-21-3p leads to an increase in the expression of E-selectin and VCAM-1 despite impaired p65 nuclear translocation. Promoters of these two genes contain binding sites for the activator protein 1 (AP-1) family of transcription factors (65, 66). Therefore, despite impaired p65 nuclear translocation in response to knockdown of miR-21-3p, IL-1 $\beta$  could act through AP-1 to induce the expression of E-selectin and VCAM-1.

#### The role of tissue specificity in the identification of miRs-21-3p and -27a-5p target genes

Recently published reports have begun to characterize gene targets of miRs-21-3p and -27a-5p in different tissue types and diseases. Activation of epidermal growth factor receptor signaling is key in the development and progression of numerous cancers such as head and neck squamous cell carcinoma. Recently published work has shown that miR-27a-5p, whose expression is decreased in head and neck squamous cell carcinoma cell lines, acts to regulate epidermal growth factor receptor signaling through the repression of v-akt murine thymoma viral oncogene homolog 1 (AKT1) and mechanistic target of rapamycin (30). NK-cells cytotoxicity is mediated through release of granules containing proteins such as perforin and

granzymes to induce apoptosis. Recent work has shown that expression of miR-27a-5p leads to loss of cytotoxicity of NK-cells through direct transcriptional repression of perforin 1 (*PRF1*) and granzyme B (*GZMB*), suggesting miR-27a-5p as repressor of apoptosis (28).

Recent work has also implicated miR-21-3p as a regulator of apoptosis and cell growth in allergic inflammation; the continued presence of eosinophils at the site of inflammation is highly dependent on the secretion of granulocyte macrophage-colony-stimulating factor (GM-CSF), which acts to suppress apoptosis (67). In culture, GM-CSF upregulates the expression of miR-21-3p, which enhances ERK activation leading to a suppression of apoptosis in eosinophils, suggesting that miR-21-3p is a key component in GM-CSF-mediated survival (27). In hepatocellular carcinoma, increased expression of *MAT2A* and *MAT2B*, which are targets of miR-21-3p, contributes to progression of the tumor growth (29). In our data, we found downregulation of the reported targets that are expressed in ECs (*AKT1*, *MAT2A*, and *MAT2B*) following 24 h overexpression (data not shown). Given the known role of NF- $\kappa$ B signaling in the promotion of cell survival and proliferation, confirmation of the downregulation of these genes in our own data is suggestive of modulation of other key processes by these two miRNAs in ECs.

In summary, our study catalogs the profile of expressed miRNAs in HAECs both pre- and postexposure to OxPAPC. Furthermore, we identified the miRNAs, miRs-21-3p and -27a-5p, to be induced in ECs in response to OxPAPC, IL-1 $\beta$  and TNFC treatment. We find these two miRNAs act to regulate NF- $\kappa$ B signaling through suppression of p65 nuclear translocation. We present evidence using the 3' UTR luciferase assay that these two miRNAs modulate the extent of activation NF- $\kappa$ B signaling by regulating genes involved in both the suppression and activation of NF- $\kappa$ B signaling. Taken together, our studies suggest an important and complex role for miRs-21-3p and -27a-5p in the regulation of NF- $\kappa$ B activation in HAECs. ■

#### REFERENCES

- Endemann, D. H., and E. L. Schiffrin. 2004. Endothelial dysfunction. *J. Am. Soc. Nephrol.* 15: 1983–1992.
- Libby, P., P. M. Ridker, and G. K. Hansson. 2009. Inflammation in atherosclerosis: from pathophysiology to practice. *J. Am. Coll. Cardiol.* 54: 2129–2138.
- Gargalovic, P. S., M. Imura, B. Zhang, N. M. Gharavi, M. J. Clark, J. Pagnon, W-P. Yang, A. He, A. Truong, S. Patel, et al. 2006. Identification of inflammatory gene modules based on variations of human endothelial cell responses to oxidized lipids. *Proc. Natl. Acad. Sci. USA.* 103: 12741–12746.
- Banfi, C., S. Colli, and S. Eligini. 2002. Oxidized LDLs influence thrombotic response and cyclooxygenase 2. *Prostaglandins Leukot. Essent. Fatty Acids.* 67: 169–173.
- Gargalovic, P. S., N. M. Gharavi, M. J. Clark, J. Pagnon, W-P. Yang, A. He, A. Truong, T. Baruch-Oren, J. A. Berliner, T. G. Kirchgesner, et al. 2006. The unfolded protein response is an important regulator of inflammatory genes in endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* 26: 2490–2496.
- Reddy, S. T., V. Grijalva, C. Ng, K. Hassan, S. Hama, R. Mottahedeh, D. J. Wadleigh, M. Navab, and A. M. Fogelman. 2002. Identification of genes induced by oxidized phospholipids in human aortic endothelial cells. *Vascul. Pharmacol.* 38: 211–218.

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

7. Yeh, M., A. L. Cole, J. Choi, Y. Liu, D. Tulchinsky, J.H. Qiao, M. C. Fishbein, A. N. Dooley, T. Hovnanian, K. Mouilleseaux, et al. 2004. Role for sterol regulatory element-binding protein in activation of endothelial cells by phospholipid oxidation products. *Circ. Res.* 95: 780–788.
8. Watson, A. D., M. Navab, S. Y. Hama, A. Sevanian, S. M. Prescott, D. M. Stafforini, T. M. McIntyre, B. N. Du, A. M. Fogelman, and J. A. Berliner. 1995. Effect of platelet activating factor-acetylhydrolase on the formation and action of minimally oxidized low density lipoprotein. *J. Clin. Invest.* 95: 774–782.
9. Zhao, B., S. A. Stachansky, R. A. Bowden, and P. D. Bowman. 2003. Effect of interleukin-1 $\beta$  and tumor necrosis factor- $\alpha$  on gene expression in human endothelial cells. *Am. J. Physiol. Cell Physiol.* 284: C1577–C1583.
10. Bartel, D. P. 2009. MicroRNAs: target recognition and regulatory functions. *Cell.* 136: 215–233.
11. Kuehnbacher, A., C. Urbich, A. M. Zeiher, and S. Dimmeler. 2007. Role of Dicer and Drosha for endothelial microRNA expression and angiogenesis. *Circ. Res.* 101: 59–68.
12. Suárez, Y., C. Fernández-Hernando, J. S. Pober, and W. C. Sessa. 2007. Dicer dependent microRNAs regulate gene expression and functions in human endothelial cells. *Circ. Res.* 100: 1164–1173.
13. Sun, H-X, D-Y. Zeng, R-T. Li, R-P. Pang, H. Yang, Y-L. Hu, Q. Zhang, Y. Jiang, L-Y. Huang, Y-B. Tang, et al. 2012. Essential role of microRNA-155 in regulating endothelium-dependent vasorelaxation by targeting endothelial nitric oxide synthase. *Hypertension.* 60: 1407–1414.
14. Marchand, A., C. Proust, P-E. Morange, A-M. Lompré, and D-A. Trégouët. 2012. miR-421 and miR-30c inhibit SERPINE 1 gene expression in human endothelial cells. *PLoS ONE.* 7: e44532.
15. Fang, Y., C. Shi, E. Manduchi, M. Civelek, and P. F. Davies. 2010. MicroRNA-10a regulation of proinflammatory phenotype in atherosusceptible endothelium in vivo and in vitro. *Proc. Natl. Acad. Sci. USA.* 107: 13450–13455.
16. Feng, B., S. Chen, K. McArthur, Y. Wu, S. Sen, Q. Ding, R. D. Feldman, and S. Chakrabarti. 2011. miR-146a-mediated extracellular matrix protein production in chronic diabetes complications. *Diabetes.* 60: 2975–2984.
17. Afonyushkin, T., O. V. Oskolkova, and V. N. Bochkov. 2012. Permissive role of miR-663 in induction of VEGF and activation of the ATF4 branch of unfolded protein response in endothelial cells by oxidized phospholipids. *Atherosclerosis.* 225: 50–55.
18. Fang, Y., and P. F. Davies. 2012. Site-specific microRNA-92a regulation of Kruppel-like factors 4 and 2 in atherosusceptible endothelium. *Arterioscler. Thromb. Vasc. Biol.* 32: 979–987.
19. Watson, A. D., N. Leitinger, M. Navab, K. F. Faul, S. Hördickö, J. L. Witztum, W. Palinski, D. Schwenke, R. G. Salomon, W. Sha, et al. 1997. Structural identification by mass spectrometry of oxidized phospholipids in minimally oxidized low density lipoprotein that induce monocyte/endothelial interactions and evidence for their presence in vivo. *J. Biol. Chem.* 272: 13597–13607.
20. Civelek, M., R. Hagopian, C. Pan, N. Che, W-P. Yang, P. S. Kayne, N. K. Saleem, H. Cederberg, J. Kuusisto, P. S. Gargalovic, et al. 2013. Genetic regulation of human adipose microRNA expression and its consequences for metabolic traits. *Hum. Mol. Genet.* 22: 3023–3037.
21. Griffiths-Jones, S., H. K. Saini, S. van Dongen, and A. J. Enright. 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36: D154–D158.
22. Shi, W., A. Oshlack, and G. K. Smyth. 2010. Optimizing the noise versus bias trade-off for Illumina whole genome expression BeadChips. *Nucleic Acids Res.* 38: e204.
23. Grant, G. R., J. Liu, and C. J. Stoeckert. 2005. A practical false discovery rate approach to identifying patterns of differential expression in microarray data. *Bioinformatics.* 21: 2684–2690.
24. Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4: 44–57.
25. Kozomara, A., and S. Griffiths-Jones. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39: D152–D157.
26. Suárez, Y., C. Wang, T. D. Manes, and J. S. Pober. 2010. Cutting edge: TNF-induced microRNAs regulate TNF-induced expression of E-selectin and intercellular adhesion molecule-1 on human endothelial cells: feedback control of inflammation. *J. Immunol.* 184: 21–25.
27. Wong, C. K., K. M. Lau, I. H. S. Chan, S. Hu, Y. Y. O. Lam, A. O. K. Choi, and C. W. K. Lam. 2013. MicroRNA-21\* regulates the pro-survival effect of GM-CSF on human eosinophils. *Immunobiology.* 218: 255–262.
28. Kim, T-D., S. U. Lee, S. Yun, H-N. Sun, S. H. Lee, J. W. Kim, H. M. Kim, S-K. Park, C. W. Lee, S. R. Yoon, et al. 2011. Human microRNA-27a\* targets Prf1 and GzmB expression to regulate NK-cell cytotoxicity. *Blood.* 118: 5476–5486.
29. Lo, T-F., W-C. Tsai, and S-T. Chen. 2013. MicroRNA-21-3p, a berberine-induced miRNA, directly down-regulates human methionine adenosyltransferases 2A and 2B and inhibits hepatoma cell growth. *PLoS ONE.* 8: e75628.
30. Wu, X., M. K. Bhayani, C. T. Dodge, M. S. Nicoloso, Y. Chen, X. Yan, M. Adachi, L. Thomas, C. E. Galcer, T. Jiffar, et al. 2013. Coordinated targeting of the EGFR signaling axis by microRNA-27a\*. *Oncotarget.* 4: 1388–1398.
31. Betel, D., M. Wilson, A. Gabow, D. S. Marks, and C. Sander. 2008. The microRNA.org resource: targets and expression. *Nucleic Acids Res.* 36: D149–D153.
32. Hoffmann, A., and D. Baltimore. 2006. Circuitry of nuclear factor kappaB signaling. *Immunol. Rev.* 210: 171–186.
33. Yeh, M., N. Leitinger, R. de Martin, N. Onai, K. Matsushima, D. K. Vora, J. A. Berliner, and S. T. Reddy. 2001. Increased transcription of IL-8 in endothelial cells is differentially regulated by TNF- $\alpha$  and oxidized phospholipids. *Arterioscler. Thromb. Vasc. Biol.* 21: 1585–1591.
34. Romanoski, C. E., N. Che, F. Yin, N. Mai, D. Poulidar, M. Civelek, C. Pan, S. Lee, L. Vakili, W-P. Yang, et al. 2011. Network for activation of human endothelial cells by oxidized phospholipids: a critical role of heme oxygenase 1. *Circ. Res.* 109: e27–e41.
35. Bochkov, V. N., O. V. Oskolkova, K. G. Birukov, A-L. Levonen, C. J. Binder, and J. Stöckl. 2010. Generation and biological activities of oxidized phospholipids. *Antioxid. Redox Signal.* 12: 1009–1059.
36. Dweep, H., C. Sticht, P. Pandey, and N. Gretz. 2011. miRWalk-database: prediction of possible miRNA binding sites by “walking” the genes of three genomes. *J. Biomed. Inform.* 44: 839–847.
37. Thomas, M., J. Lieberman, and A. Lal. 2010. Desperately seeking microRNA targets. *Nat. Struct. Mol. Biol.* 17: 1169–1174.
38. Lee, S., K. G. Birukov, C. E. Romanoski, J. R. Springstead, A. J. Lusis, and J. A. Berliner. 2012. Role of phospholipid oxidation products in atherosclerosis. *Circ. Res.* 111: 778–799.
39. McCall, M. N., O. A. Kent, J. Yu, K. Fox-Talbot, A. L. Zaiman, and M. K. Halushka. 2011. MicroRNA profiling of diverse endothelial cell types. *BMC Med. Genomics.* 4: 78.
40. Pritchard, C. C., H. H. Cheng, and M. Tewari. 2012. MicroRNA profiling: approaches and considerations. *Nat. Rev. Genet.* 13: 358–369.
41. Voellenkle, C., J. Van Rooij, A. Guffanti, E. Brimi, P. Fasanaro, E. Isaia, L. Croft, M. David, M. C. Capogrossi, A. Moles, et al. 2012. Deep-sequencing of endothelial cells exposed to hypoxia reveals the complexity of known and novel microRNAs. *RNA.* 18: 472–484.
42. Pedrioli, D. M. L., T. Karpanen, V. Dabouras, G. Jurisic, G. van de Hock, J. W. Shin, D. Marino, R. E. Kälin, S. Leidel, P. Cincelli, et al. 2010. miR-31 functions as a negative regulator of lymphatic vascular lineage-specific differentiation in vitro and vascular development in vivo. *Mol. Cell. Biol.* 30: 3620–3634.
43. Sun, X., B. Icli, A. K. Wara, N. Belkin, S. He, L. Kobzik, G. M. Hunninghake, M. P. Vera, M. Registry, T. S. Blackwell, et al. 2012. MicroRNA-181b regulates NF- $\kappa$ B-mediated vascular inflammation. *J. Clin. Invest.* 122: 1973–1990.
44. Wang, S., A. B. Aurora, B. A. Johnson, X. Qi, J. McAnally, J. A. Hill, J. A. Richardson, R. Bassel-Duby, and E. N. Olson. 2008. The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. *Dev. Cell.* 15: 261–271.
45. Schober, A., M. Nazari-Jahantigh, Y. Wei, K. Bidzhekov, F. Gremse, J. Grommes, R. T. A. McGens, K. Heyil, H. Noels, M. Hristov, et al. 2014. MicroRNA-126-5p promotes endothelial proliferation and limits atherosclerosis by suppressing Dlk1. *Nat. Med.* 20: 368–376.
46. Urbich, C., D. Kaluza, T. Frömel, A. Knau, K. Bennewitz, R. A. Boon, A. Bonauer, C. Doebele, J-N. Boeckel, E. Hergenreider, et al. 2012. MicroRNA-27a/b controls endothelial cell repulsion and angiogenesis by targeting semaphorin 6A. *Blood.* 119: 1607–1616.
47. Bridge, G., R. Monteiro, S. Henderson, V. Emuss, D. Lagos, D. Georgopoulou, R. Patient, and C. Boshoff. 2012. The microRNA-30 family targets DLL4 to modulate endothelial cell behavior during angiogenesis. *Blood.* 120: 5063–5072.
48. Fiedler, J. V. Jazbutyte, B. C. Kirchmaier, S. K. Gupta, J. Lorenzen, D. Hartmann, P. Galuppo, S. Kneitz, J. T. G. Pena, C. Sohn-Lee, et al. 2011. MicroRNA-24 regulates vascularity after myocardial infarction. *Circulation.* 124: 720–730.

Supplemental Material can be found at:  
<http://www.jlr.org/content/suppl/2014/10/19/jlr.M052670.DC1.html>

49. Harris, T. A., M. Yamakuchi, M. Ferlito, J. T. Mendell, and C. J. Lowenstein. 2008. MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1. *Proc. Natl. Acad. Sci. USA.* 105: 1516–1521.
50. Mohan, S., N. Mohan, and E. A. Sprague. 1997. Differential endothelial activation of NF-kappa B in human aortic cells conditioned to specific flow environments. *Am. J. Physiol. Cell Physiol.* 273: C572–C578.
51. True, A. L., A. Rahman, and A. B. Malik. 2000. Activation of NF-kappaB induced by H(2)O(2) and TNF-alpha and its effects on ICAM-1 expression in endothelial cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 279: L302–L311.
52. Faure, E., O. Equils, P. A. Sieling, L. Thomas, F. X. Zhang, C. J. Kirschning, N. Polentarutti, M. Muzio, and M. Arditi. 2000. Bacterial lipopolysaccharide activates NF-kappa B through Toll-like receptor 4 (TLR-4) in cultured human dermal endothelial cells. Differential expression of TLR-4 and TLR-2 in endothelial cells. *J. Biol. Chem.* 275: 11058–11063.
53. Tili, E., J. J. Michaille, A. Cimino, S. Costinean, C. D. Dumitru, B. Adair, M. Fabbri, H. Alder, C. G. Liu, G. A. Calin, et al. 2007. Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-alpha stimulation and their possible roles in regulating the response to endotoxin shock. *J. Immunol.* 179: 5082–5089.
54. Chen, R., A. B. Alvero, D. A. Silasi, M. G. Kelly, S. Fest, I. Visintin, A. Leiser, P. E. Schwartz, T. Rutherford, and G. Mor. 2008. Regulation of IKKbeta by miR-199a affects NF-kappaB activity in ovarian cancer cells. *Oncogene* 27: 4712–4723.
55. Li, T., M. J. Morgan, S. Choksi, Y. Zhang, Y-S. Kim, and Z. Liu. 2010. MicroRNAs modulate the noncanonical transcription factor NF-kappaB pathway by regulating expression of the kinase IKKalpha during macrophage differentiation. *Nat. Immunol.* 11: 799–805.
56. Min, H., and S. Yoon. 2010. Got target? Computational methods for microRNA target prediction and their extension. *Exp. Mol. Med.* 42: 233–244.
57. Hausser, J., and M. Zavolan. 2014. Identification and consequences of miRNA-target interactions - beyond repression of gene expression. *Nat. Rev. Genet.* 15: 599–612.
58. Sethupathy, P., M. Megraw, and A. Hatzigeorgiou. 2006. A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat. Methods.* 3: 881–886.
59. Alexiou, P., M. Maragkakis, G. L. Papadopoulos, M. Reczko, and A.G. Hatzigeorgiou. 2009. Lost in translation: an assessment and perspective for computational microRNA target identification. *Bioinformatics* 25: 3049–3055.
60. Martin, H. C., S. Wani, A. L. Steptoe, K. Krishnan, K. Nones, E. Nourbakhsh, A. Vlassov, S. M. Grimmond, and N. Cloonan. 2014. Imperfect centered miRNA binding sites are common and can mediate repression of target mRNAs. *Genome Biol.* 15: R51.
61. Ebert, M. S., and P. A. Sharp. 2012. Roles for microRNAs in conferring robustness to biological processes. *Cell.* 149: 515–524.
62. Choi, W-Y., A. J. Giraldez, and A. F. Schier. 2007. Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. *Science.* 318: 271–274.
63. Mauro, C., F. Pacifico, A. Lavorgna, S. Mellone, A. Iannetti, R. Acquaviva, S. Formisano, P. Vito, and A. Leonardi. 2006. ABIN-1 binds to NEMO/IKKgamma and co-operates with A20 in inhibiting NF-kappaB. *J. Biol. Chem.* 281: 18482–18488.
64. O'Neill, L. A., and C. Greene. 1998. Signal transduction pathways activated by the IL-1 receptor family: ancient signaling machinery in mammals, insects, and plants. *J. Leukoc. Biol.* 63: 650–657.
65. Jensen, L. E., and A. S. Whitehead. 2003. ELAM-1/E-selectin promoter contains an inducible AP-1/CREB site and is not NF-kappa B-specific. *Biotechniques.* 35: 54–56, 58.
66. Cybulsky, M. I., J. W. Fries, A. J. Williams, P. Sultan, R. Eddy, M. Byers, T. Shows, M. A. Gimbrone, and T. Collins. 1991. Gene structure, chromosomal location, and basis for alternative mRNA splicing of the human VCAM1 gene. *Proc. Natl. Acad. Sci. USA.* 88: 7859–7863.
67. Simon, H. U., S. Yousefi, B. Dibbert, F. Levi-Schaffer, and K. Blaser. 1997. Anti-apoptotic signals of granulocyte-macrophage colony-stimulating factor are transduced via Jak2 tyrosine kinase in eosinophils. *Eur. J. Immunol.* 27: 3536–3539.

## **Chapter 3**

**Mapping genetic contributions to cardiac pathology induced by Beta-adrenergic stimulation in mice.**

**Chapter 3 Preface:** This chapter is a reprint of a manuscript describing the identification and *in vivo* validation of the gene *Abcc6* as a novel contributor to the development of cardiac fibrosis *in vivo*. As a collaborator in this study my major contribution to the manuscript was the identification of candidate fibrosis genes using systematic analysis of the 6 isoproterenol induced fibrosis loci using a combination of tissue expression, correlation analysis, eQTL analysis, non-synonymous variant identification and functional variant effect analysis. Furthermore, I assisted in the preparation and collection of all samples related to the *in vivo* validation of *Abcc6* in addition to assisting in the preparation of the manuscript.

## Mapping Genetic Contributions to Cardiac Pathology Induced by Beta-Adrenergic Stimulation in Mice

Christoph D. Rau, PhD\*; Jessica Wang, MD, PhD\*; Rozeta Avetisyan, BS; Milagros C. Romay, BS; Lisa Martin, PhD; Shuxun Ren, MD; Yibin Wang, PhD; Aldons J. Lusis, PhD

**Background**—Chronic stress-induced cardiac pathology exhibits both a wide range in severity and a high degree of heterogeneity in clinical manifestation in human patients. This variability is contributed to by complex genetic and environmental etiologies within the human population. Genetic approaches to elucidate the genetics underlying the acquired forms of cardiomyopathies, including genome-wide association studies, have been largely unsuccessful, resulting in limited knowledge as to the contribution of genetic variations for this important disease.

**Methods and Results**—Using the  $\beta$ -adrenergic agonist isoproterenol as a specific pathological stressor to circumvent the problem of etiologic heterogeneity, we performed a genome-wide association study for genes influencing cardiac hypertrophy and fibrosis in a large panel of inbred mice. Our analyses revealed 7 significant loci and 17 suggestive loci, containing an average of 14 genes, affecting cardiac hypertrophy, fibrosis, and surrogate traits relevant to heart failure. Several loci contained candidate genes which are known to contribute to Mendelian cardiomyopathies in humans or have established roles in cardiac pathology based on molecular or genetic studies in mouse models. In particular, we identify *Abcc6* as a novel gene underlying a fibrosis locus by validating that an allele with a splice mutation of *Abcc6* dramatically and rapidly promotes isoproterenol-induced cardiac fibrosis.

**Conclusions**—Genetic variants significantly contribute to the phenotypic heterogeneity of stress-induced cardiomyopathy. Systems genetics is an effective approach to identify genes and pathways underlying the specific pathological features of cardiomyopathies. *Abcc6* is a previously unrecognized player in the development of stress-induced cardiac fibrosis. (*Circ Cardiovasc Genet.* 2015;8:40-49. DOI: 10.1161/CIRCGENETICS.113.000732.)

**Key Words:** catecholamine ■ genome-wide association scan ■ genomics ■ heart failure ■ mouse

Heart failure (HF) is a common cause of death with a lifetime risk of  $\geq 1$  in 9 for both men and women in developed countries.<sup>1</sup> Heart failure is a complicated syndrome, characterized by a large number of pathological changes, such as contractile dysfunction, cardiomyocyte hypertrophy, edema, and myocardial fibrosis.<sup>2-4</sup> The onset and severity of these pathological manifestations are highly heterogeneous among HF patients, likely because of complex interactions between the genetic variants and the pathological stressors, including mechanical overload and humoral overstimulation. Indeed, several humoral factors, such as catecholamines and angiotensin II, are known to play key roles in triggering HF; however, the genetic variations underlying the pathological outcome in response to these stressors remain elusive. Dissecting the genetic contributions to specific pathological changes in the failing heart would provide important insights for the future development of personalized diagnoses and targeted therapies.

### Clinical Perspective on p 49

In contrast to many other common disorders, genome-wide association studies (GWAS) of HF have had modest success in elucidating the genetics underlying this complex disease. Only 2 heart failure-related loci<sup>5</sup> have reached accepted levels of genome-wide significance, despite meta-analyses of tens of thousands of patients.<sup>6,7</sup> The challenge of performing GWAS in human HF is likely because of the complex nature of the disease, which can arise as a result of multiple underlying etiologies, such as myocardial infarction, hypertension, or metabolic disorders, each of which are complex traits with significant environmental confounders.<sup>1</sup> Attempts to dissect the genetics of HF traits in rodents have been only modestly successful; although several loci for hypertrophy and fibrosis have been identified, the poor mapping resolution of traditional linkage analyses has complicated the identification of the underlying genes.<sup>8-11</sup> The development of a method to perform high resolution, association-based mapping of complex

Received December 19, 2013; accepted October 28, 2014.

From the Department of Microbiology, Immunology and Molecular Genetics, Department of Human Genetics (C.D.R., R.A., M.R., A.J.L.), Department of Medicine, Division of Cardiology, David Geffen School of Medicine (J.W., L.M., A.J.L.), and Departments of Anesthesiology, Physiology, and Medicine, Cardiovascular Research Laboratories, David Geffen School of Medicine (S.R., Y.W.), University of California, Los Angeles, Los Angeles, CA.

\*Drs Rau and J. Wang contributed equally to this work.

The Data Supplement is available at <http://circgenetics.ahajournals.org/lookup/suppl/doi:10.1161/CIRCGENETICS.113.000732/-/DC1>.

Correspondence to Aldons J. Lusis, PhD, Department of Medicine/Division of Cardiology, A2-237 Center for the Health Sciences, UCLA School of Medicine, University of California, Los Angeles, CA 90095. E-mail [jlusis@mednet.ucla.edu](mailto:jlusis@mednet.ucla.edu) or Yibin Wang, PhD, Department of Anesthesiology, Division of Molecular Medicine, University of California, Los Angeles, BII-569 CHS, Box 957115, Los Angeles, CA 90095. E-mail [yibinwang@mednet.ucla.edu](mailto:yibinwang@mednet.ucla.edu)

© 2014 American Heart Association, Inc.

*Circ Cardiovasc Genet* is available at <http://circgenetics.ahajournals.org>

DOI: 10.1161/CIRCGENETICS.113.000732

Downloaded from <http://circgenetics.ahajournals.org/> at CONS CALIFORNIA DIG LIB on May 9, 2016

traits in mice<sup>12</sup> provided an opportunity to identify genetic factors contributing to common forms of HF under defined stress conditions.

In this study, we have conducted a comprehensive phenotypic characterization in a large panel of densely genotyped inbred mice from the hybrid mouse diversity panel (HMDP)<sup>12</sup> following chronic treatment with a  $\beta$ -adrenergic agonist, isoproterenol (ISO). A wide spectrum of phenotypic changes was observed among the HMDP mice in ISO-induced cardiac hypertrophy, fibrosis, and peripheral edema. Using GWAS, we uncovered 7 significant loci and 17 suggestive loci, each containing an average of 14 genes. Several of these loci included genes with established causal roles in familial cardiomyopathies in humans or heart failure phenotypes in experimental models. In addition, we identified *Abcc6* as a previously unrecognized regulator of ISO-induced cardiac fibrosis. Therefore, our study provides clear evidence that genetic variants have a significant contribution to the phenotypic heterogeneity of stress-induced cardiomyopathy.

## Materials and Methods

### Ethics Statement

All animal experiments were conducted following guidelines established and approved by the University of California, Los Angeles Institutional Animal Care and Use Committee. All surgery and echocardiography was performed under isoflurane anesthesia, and every effort was made to minimize suffering.

### Online Database

All results and data can be accessed at <http://systems.genetics.ucla.edu/data>.

### Mice and Isoproterenol Treatment

The mouse strains listed in Table 1 in the Data Supplement were obtained from The Jackson Laboratory and then bred in our colony. All mice have been previously genotyped at over 130 000 locations. ISO (30 mg per kg body weight per day, Sigma) was administered for 21 days in 8- to 10-week-old female mice using ALZET osmotic minipumps, which were surgically implanted intraperitoneally.

*Abcc6* knockout (KO) and transgenic mice<sup>13,14</sup> underwent the same protocol as described above, although both male and female mice were used in the analysis. No significant difference between genders was observed as a result of ISO treatment in these KO and transgenic animals.

### Heart Weights

At day 21, mice were euthanized and body weight recorded. The heart was removed and weighed, then separated into its 4 component chambers, each of which was individually weighed as well. Each chamber of the heart was immediately frozen in liquid nitrogen for any future analysis and stored in a  $-80^{\circ}$  freezer. Lung and liver were removed and weighed. Additionally, the adrenal glands were removed, weighed, and frozen in liquid nitrogen.

### Fibrosis and Calcification

A portion of the left ventricle (LV) was placed in formalin for  $\geq 48$  hours for preservation of ultrastructure. These samples were then washed with distilled water and sent to UCLA Department of Pathology and Laboratory Medicine for paraffin embedding and staining using Masson's Trichrome for fibrosis and Alizarin Red for calcification. Sections were analyzed using a Nikon Eclipse, TE2000-U microscope and images captured of the entire cross-section of the heart. Fibrosis was quantified using the Nikon Imagine

System Elements AR program by comparing the amount of tissue stained blue (for collagen) or red (for calcification) to the total tissue area. To confirm our results, we examined a subset of the strains using Sirius Red, another fibrosis-marking stain, and observed high concordance between our samples ( $R=0.75$ , data not shown). As expected,<sup>4,15</sup> we observed strong correlations between cardiac fibrosis and total heart weight ( $P=1.1E-07$ ). Our results compare favorably to prior quantifications of fibrosis in a limited number of strains.<sup>16</sup>

Mice used for the *Abcc6* validation experiments underwent an identical protocol to the one described above for Masson Trichrome and Alizarin Red staining, using 5 sections per heart for both Trichrome and Alizarin Red staining.

### Echocardiography

Transthoracic echocardiograms were performed using the Vevo 770 ultrasound system (VisualSonics, Inc., Toronto, ON, Canada). Inhaled isoflurane (1.25% during induction and 1% during maintenance) was administered to ensure adequate sedation in the meantime maintaining heart rate  $>450$  beats per minute. A parasternal long-axis B-mode image was obtained. The maximal long-axis of the LV was positioned perpendicular to the ultrasound beam. A  $90^{\circ}$  rotation of the ultrasound probe at the papillary muscle level was performed to obtain a parasternal short-axis view of the LV. A M-mode image was captured to document LV dimensions. Then a semiapical long-axis view of the LV was obtained. The LV ejection time, E and A wave velocities were obtained from this view using pulse wave Doppler. Images were saved for analysis at a later time point using the Vevo 770 cardiac analysis package. In summary, a baseline echocardiogram was performed on all of the mice. Among control mice, a second echocardiogram was performed in 70 mouse strains at week 3. In ISO-treated mice, serial echocardiograms were performed at 1, 2, and 3 weeks. A single operator, who followed a standard operating protocol detailed above, performed all of the echocardiograms. Saved images were analyzed at a later time point by a single observer who was blinded to mouse strains.

### Association Analysis

Unless otherwise noted, all analyses were performed using the R software environment. We performed the association testing of each single-nucleotide polymorphism (SNP) with a linear mixed model, which accounts for the population structure among the  $n$  animals using the following model<sup>17</sup>:

$$y = \mathbf{1}_n + m + xb + u + e$$

where  $m$  is the mean,  $b$  is the allele effect of the SNP,  $x$  is the  $(n \times 1)$  vector of observed genotypes of the SNP (using additive coding of 0,0.5,1),  $u$  is the random effects because of genetic relatedness with  $\text{var}(u) = \sigma_u^2 K$ , and  $e$  is the random noise with  $\text{var}(e) = \sigma_e^2 I$ .  $K$  denotes the identity-by-state kinship matrix estimated from all the SNPs,  $I$  denotes the  $(n \times n)$  identity matrix, and  $\mathbf{1}_n$  is the  $(n \times 1)$  vector of ones. We estimated  $\sigma_u^2$  and  $\sigma_e^2$  using restricted maximum likelihood and computed  $P$  values using the standard  $F$  test to test the null hypothesis  $b=0$ . All phenotypes were examined for deviation from normality by visual examination of a normal probability plot. The right ventricular (RV) ratio phenotype was found to deviate somewhat from normality; however, to keep the analysis consistent across phenotypes, no transformation of the data was performed. Prior work<sup>17</sup> has demonstrated that for the HMDP, 4.1E-6 is the correct significance threshold for a single trait. We conservatively estimated that we were observing 10 independent phenotypes in our data and determined our final significance threshold, 4.1E-7, by Bonferroni correction. Linkage disequilibrium (LD) was determined by calculated pairwise  $r^2$  SNP correlations for each chromosome. Approximate LD boundaries were determined by visualizing  $r^2 > 0.8$  correlations in MATLAB (MathWorks).

### Locus Overlap with Other Studies

Gwas.gov was queried for all human GWAS loci for the terms heart failure or cardiac hypertrophy. All loci with  $P$  value  $< 5E-7$  were

selected. The NCBI homology maps (<http://www.ncbi.nlm.nih.gov/projects/homology/maps/>) were used to find syntenic location of the 5 Mb region surrounding the peak SNP of the human HF locus. If a mapped syntenic region overlapped with the LD block of a suggestive locus from our study, it was considered a positive hit. A significance  $P$  value was obtained by permutation testing, in which all suggestive loci were randomly placed across the genome and the number of overlaps measured a total of 100 000 times. Final significance was calculated as the number of permutations which surpassed the observed number of overlaps.

### Microarray and eQTL Analysis

Following homogenization of LV tissue samples in QIAzol, RNA was extracted using the Qiagen miRNAeasy extraction kit and verified as having a RIN > 7 by Agilent Bioanalyzer. Two RNA samples were pooled for each strain/experimental condition, whenever possible, and arrayed on Illumina Mouse Reference 8 version 2.0 chips. Analysis was conducted using the Neqc algorithm included in the limma R package,<sup>18</sup> and batch effects addressed through the use of COMBAT.<sup>19</sup> Expression quantitative trait loci (eQTLs) were then calculated for 13 155 expressed genes using EMMA, as described above. Significance thresholds were calculated as in Parks et al.<sup>20</sup> Briefly, *cis*-eQTLs were calculated using a false discovery rate (FDR) of 5% for all SNPs that lay within 1 Mb of any probe (roughly 100 SNPs), using standard permutation analysis methods (total of 100 permutations of all data), previously used for *cis*-eQTL analysis.<sup>21–23</sup> Our determined cutoff of 3.6E-3 further takes into account LD, which further reduces the effective number of SNPs in each 2 Mb window surrounding the peak SNP. We have previously observed that *cis*-eQTL at this level are highly conserved in the HMDP.<sup>24</sup> *Trans*-eQTLs were calculated using the overall HMDP cutoff as determined in Kang et al and described above.<sup>17</sup>

## Results

### Pathological Analysis of ISO-Induced Cardiomyopathy in HMDP Mice

$\beta$ -adrenergic stimulation is considered a common and critical driving force behind ongoing hypertrophy and progression to heart failure.<sup>25</sup> We treated mice chronically with ISO, a synthetic nonselective  $\beta$ -adrenergic agonist.<sup>26,27</sup> 748 mice from 105 different strains of the HMDP were divided into control (average 2.2 per strain) and treated (average 4.1 per strain) cohorts (Table I in the Data Supplement). Treated mice were implanted with an Alzet micropump and given 30 mg/kg/d of ISO for 3 weeks, at which point all mice were euthanized. We characterized a variety of phenotypes to capture specific portions of the complex heart failure syndrome. For this report, cardiac hypertrophy and pulmonary and liver edema were assessed by measuring the weights of the 4 cardiac chambers, the lungs, and the liver. Cardiac fibrosis, a phenotype which is difficult to study in humans, was measured by histological quantification of fibrotic tissue area as a percentage of all tissue area in LV sections stained using Masson Trichrome. Functional analysis of the mouse hearts were performed using echocardiography. We observed statistically significant correlations between our calculated LV weights (LVW) from echocardiography and our measured LVW after harvesting ( $R=0.82$ ,  $P=5E-24$ ) as well as between LVW and LV internal dimension ( $R=0.26$ ,  $P=1.3E-11$ ). Further analysis of the functional data, including association analyses for each observed functional phenotype, are still under preparation and will be reported in a subsequent article.

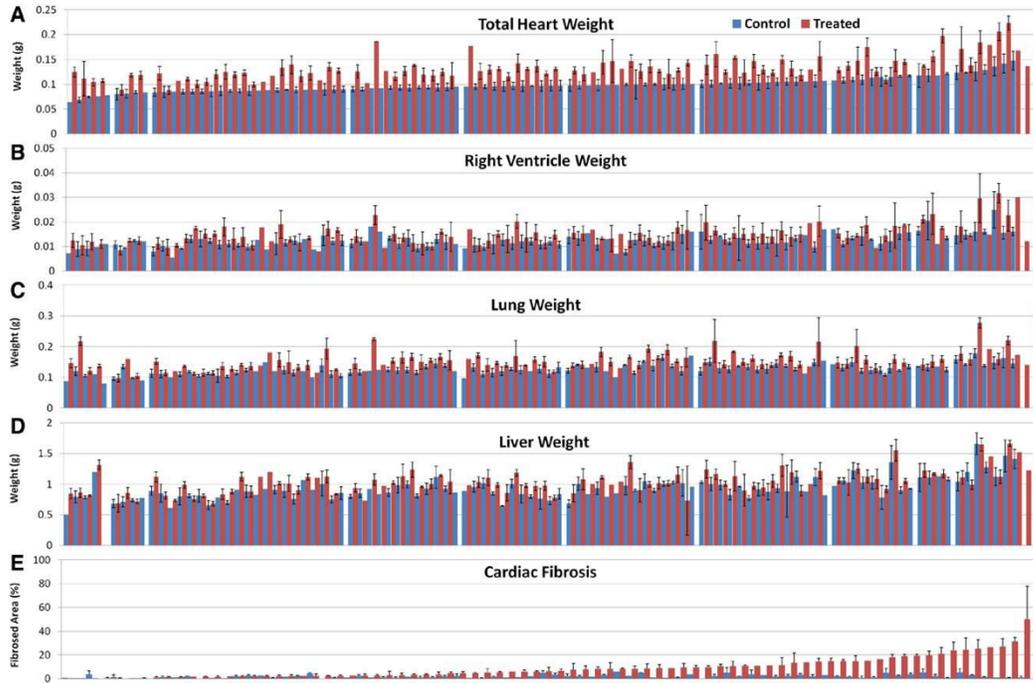
As shown in Figure 1, we observed striking differences in cardiac hypertrophy, fibrosis, and degrees of pulmonary

and hepatic edema among the strains. Our results are consistent with another report from a more limited strain survey.<sup>28</sup> After treatment with ISO, mice were slightly heavier than their paired controls (23.5 g versus 21.83 g;  $P=0.009$ ), likely because of edema. More significant changes in weight were observed for our traits of interest, such as total heart weight ( $P=5.3E-24$ ) and lung weight (4.6E-12; Table II in the Data Supplement). Although it is likely that the ISO-treated organ weights are influenced by the slight increase in body weight, the magnitude of the results we observe in our study leads us to conclude that this influence is minimal when compared with the direct effects of ISO stimulation on the heart itself.

Of the 470 mice assigned to the treatment cohort, 139 (29.6%) died before the end of the protocol, most (127) within the first 48 hours of treatment, whereas none of the control cohort died (Figure 1 in the Data Supplement). There was no observable significant differences in baseline conditions between mice which survived the initial ISO challenge and mice which died after ISO challenge. Furthermore, we did not observe significant correlation between the strains which showed high mortality before end of protocol and any of our expressed genes before or after treatment (the most significant correlation was for the Riken gene 8430432M10RIK [ $R^2=0.16$ ,  $P=5.6E-5$ ], which fails to satisfy a Bonferroni-corrected threshold of 3E-6). Among the phenotypic traits (Bonferroni-corrected threshold of 4.8E-4), we observed that the post-ISO Mean Normalized Systolic Ejection Rate functional trait demonstrated significant correlation ( $R^2=0.14$ ,  $P=3E-4$ ) with premature mortality. This relationship will be explored in greater detail in our article focusing on echocardiographic functional traits. GWA on this premature death trait revealed a single locus (peak SNP rs29166005,  $P=9E-6$ ), which contributes suggestively to this phenotype. These results suggest that the cause of our observed ISO-induced death is likely linked to underlying genetic effects or the interaction between genes and the ISO treatment.

### Genomewide Association

Association analysis was performed using  $\approx 132$  000 SNPs across the genome with the EMMA algorithm<sup>17</sup> to correct for population structure. In addition to the absolute tissue weight measurements, analyses were performed on the ratios of each treated weight to its corresponding control weight as a measure of responsiveness to ISO treatment. Prior work with EMMA and the HMDP, using simulation and permutation, has suggested that an appropriate genome-wide significance threshold for a single trait is 4.1E-06.<sup>12</sup> This is approximately equivalent to a Bonferroni correction.<sup>12</sup> To correct for multiple comparisons, we have chosen the threshold of 4.1E-07 and a minimum minor allele frequency of 7.5% for our study. Given that the traits are correlated, this threshold (10-fold lower than the genome-wide significance level for a single trait) is conservative. Using these thresholds, we have identified 7 significant loci and 17 additional loci which matched the nominal significance threshold of 4.1E-06 (Tables 1 and 2; Table III in the Data Supplement). Although linkage analysis in mice typically exhibits a resolution of tens of Mb,<sup>8,29</sup> the loci identified in this study averaged 1 to 2 Mb in size, based on LD, with the majority being <1 Mb.



**Figure 1.** Wide variation in heart failure (HF) traits between the strains of the hybrid mouse diversity panel (HMDP). **A**, Total heart weight; **B**, Right ventricular weight; **C**, Lung weight; **D**, Liver weight; and **E**, Cardiac fibrosis. **A–D**, Organized by the untreated heart weight of the strain and display mean±standard deviation. For enlarged figures with strain names, see Figures III–VII in the Data Supplement.

The RV weight (RVW) and LVW variations mirrored each other closely, with associations being somewhat stronger for RVW (Figure 1B), although each locus identified in the RV was also detected at a lower level of significance in the LV

(Figure II in the Data Supplement). In total, we observed 3 significant and 5 suggestive loci corresponding to treated RVW (Figure 2A), 1 significant and 1 suggestive locus for the ratio of treated to untreated RVW (Figure 2B), and a single

**Table 1. Significant Heart Failure Trait Loci Identified in HMDP GWAS**

Phenotype	Chr	Bp	P Value	LD	N	Gene
<b>Hypertrophic loci</b>						
RV	5	137934905	3.49E-10	137.93–138.15	11	<i>Mospd3</i>
RV	9	40202022	8.41E-08	39.77–40.52	15	<i>Scn3b</i>
RV	10	49818583	2.80E-07	48.19–54.24	22	<i>Pln</i>
RV ratio	9	80542295	2.94E-07	80.00–80.99	2	<i>Myo6</i>
<b>Fluid retention loci</b>						
Liver	7	15251391	1.93E-07	15.13–18.75	57	<i>Calm3</i>
Lung	6	53975816	2.90E-07	53.88–55.57	17	<i>Aqp1</i>
<b>Isoproterenol treated fibrosis loci</b>						
Fibrosis	X	10277028	4.10E-07	5–12.5	127	<i>Spx</i>

The significance threshold is defined as *P* value <4.1E-07. RV, Liver, and Lung represents isoproterenol-treated right ventricular, liver, and lung weights at week 3, respectively. Fibrosis represents isoproterenol-treated LV fibrosis at week 3. For each locus, the peak single-nucleotide polymorphism location, given by chromosome (Chr) and base pair position (Bp) in the NCBI-build-37 assembly, and association *P* value are reported, along with the number of genes (*N*) within the estimated LD block (LD) surrounding the peak single-nucleotide polymorphism and the top candidate gene (Gene). Bold entries represent genes which contain nonsynonymous mutations within the HMDP as reported by the Wellcome Trust Mouse Genome Project, whereas underlined entries possess significant cis-eQTLs.

eQTLs indicates expression quantitative trait loci; HMDP, hybrid mouse diversity panel; LD, linkage disequilibrium; and RV, right ventricle.

**Table 2. Suggestive Heart Failure Trait Loci Identified in HMDP GWA**

Phenotype	Chr	Bp	P Value	LD	N	Gene
<b>Hypertrophic loci</b>						
RV	1	134467906	5.75E-07	133.78–134.53	14	...
RV	5	23873494	1.23E-06	23.82–24.47	20	<i>Prkag2</i>
RV	11	47181489	2.15E-06	46.18–49.3	41	<i>Sgcd</i>
RV	3	136305887	7.83E-07	136.04–136.79	1	<i>Ppp3ca</i>
<b>Ratio</b>						
RA Ratio	7	142011844	1.41E-06	141.50–144.81	15	<i>Mgmt</i>
<b>Fluid retention loci</b>						
Liver	10	49468021	3.48E-06	48.19–54.24	22	<i>Pln</i>
Lung	5	111867706	1.28E-06	110.87–112.87	22	<i>Miat</i>
Lung	7	81841621	3.88E-06	79.8–82.2	6	<i>Slco3a1</i>
Lung	14	14941056	3.34E-06	8.5–21.5	50	<i>Iqgap1</i>
Lung	19	27061190	2.01E-06	26.68–27.43	1	<i>Ppp3cb</i>
Lung	19	27061190	2.01E-06	26.68–27.43	1	<i>Vldlr</i>
<b>Baseline fibrosis loci</b>						
Fibrosis	2	139163425	2.51E-06	13.7–14.0	6	<i>Jag1</i>
Fibrosis	4	84420058	2.20E-06	84–85	2	<i>Cntln</i>
Fibrosis	7	73365047	1.31E-06	72.3–74.3	7	<i>Tjp1</i>
<b>Isoproterenol-treated fibrosis loci</b>						
Fibrosis	7	52946331	7.11E-07	52.85–53.42	28	<i>Abcc6</i>
Fibrosis	7	68593223	1.40E-06	60.5–69.5	18	<i>Snrpn</i>
Fibrosis	7	73365047	1.40E-06	72.3–74.3	7	<i>Tjp1</i>
Fibrosis	15	69907056	9.60E-07	68.4–71.4	3	<i>Col22a1</i>

The suggestive threshold is defined as  $P$  value  $<4.1E-06$ . RV, liver, and lung represents isoproterenol-treated right ventricular, liver, and lung weights at week 3, respectively. RV and RA ratio represents the ratio of isoproterenol-treated versus control right ventricle or right atrium weight at week 3. Fibrosis represents either control or isoproterenol-treated LV fibrosis at week 3. For each locus, the peak single-nucleotide polymorphism location, given by chromosome (Chr) and base pair position (Bp) in the NCBI-build-37 assembly, and association  $P$  value are reported, along with the number of genes (N) within the estimated LD block (LD) surrounding the peak single-nucleotide polymorphism and the top candidate gene (Gene). Bold entries represent genes which contain nonsynonymous mutations within the HMDP as reported by the Wellcome Trust Mouse Genome Project, whereas underlined entries possess significant *cis*-eQTLs.

eQTLs indicates expression quantitative trait loci; HMDP, hybrid mouse diversity panel; LD, linkage disequilibrium; and RV, right ventricle.

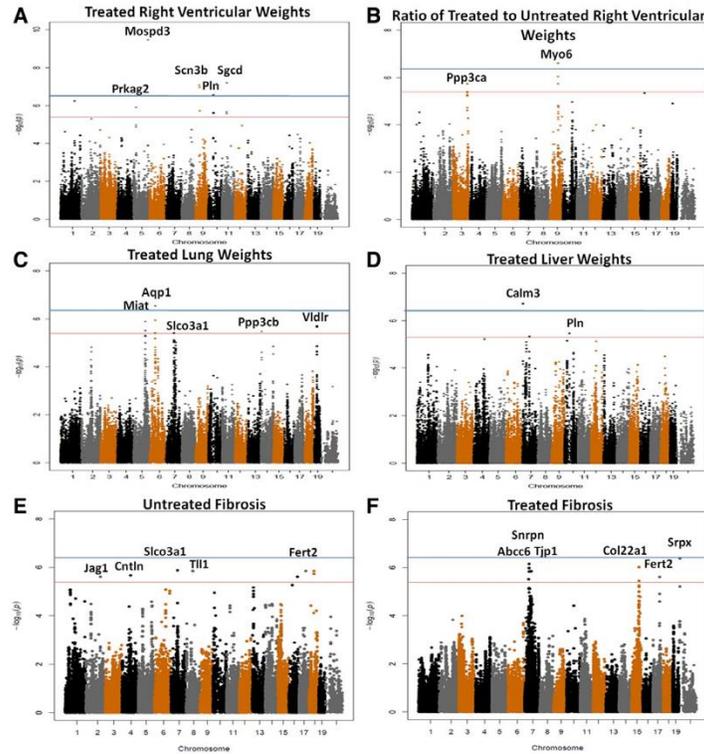
suggestive locus for the ratio of right atrial weight. Similar to the heart weights, we observed marked variation of liver and lung weights after ISO treatment across the HMDP (Figure 1C and 1D). Lung weight in particular showed a robust increase with ISO treatment. We observed 1 significant and 4 suggestive loci corresponding to ISO-treated lung weights (Figure 2C) and 1 significant and 1 suggestive locus corresponding to ISO-treated liver weights (Figure 2D). Cardiac fibrosis also varied significantly in both baseline and treated mice, with the extent of fibrosis being much greater in treated mice (Figure 1E). We observed a total of 3 suggestive loci for cardiac fibrosis in untreated animals (Figure 2E) and 1 significant and 4 suggestive loci in treated animals (Figure 2F).

#### eQTL Analysis for Candidate Genes From ISO-Treated HMDP Mice

To help identify candidate genes at the heart failure associated loci, we performed global expression analysis of LV heart tissue from 92 strains of ISO-treated mice. The loci controlling gene expression levels were mapped using EMMA and are referred to as eQTL. eQTLs were termed *cis* if the locus maps

within 1 Mb of the gene and otherwise were termed *trans*. Overall, we observed 3093 *cis* eQTL (False Discovery Rate  $5\% = P < 3.6E-3$ , in line with previous measures in the HMDP).<sup>20</sup> Additionally, the Wellcome Trust Mouse Genomes Project sequencing database,<sup>30</sup> which has the full genomic sequence of 10 strains in our panel, was used to examine genomic variations, such as missense, nonsense, or splicing variations, in each locus. Together, these 2 approaches provided a powerful and systematic method for the identification of causal genes within each locus. All significant and suggestive loci as well as gene expression data are available at <http://systems.genetics.ucla.edu/data>.

Using eQTL analysis combined with GWAS, we identified a causal gene in one of the cardiac hypertrophy loci on chromosome 3 (Figure 3). The peak SNP ( $P = 1.9E-6$ ), for the trait of treated-to-untreated right RVW ratio, lies between the second and third exons of *Ppp3ca*, encoding the alpha isozyme of calcineurin A, which is also the only gene contained within the LD block surrounding the significantly associated SNPs. Calcineurin A is a known target of  $\beta$ -adrenergic signaling, with a well-described role in ISO-induced hypertrophy.<sup>26</sup>



**Figure 2.** Manhattan plots of heart failure (HF) traits. **A**, Treated right ventricular weights; **B**, ratio of treated to untreated right ventricular weights; **C**, treated lung weights; **D**, treated liver weights; **E**, baseline cardiac fibrosis; and **F**, treated cardiac fibrosis. The red line indicates the threshold for suggestive association ( $4.2E-6$ ) between a single-nucleotide polymorphism and a phenotype, whereas the blue line indicates the threshold for significant association ( $4.2E-7$ ). Proposed candidate genes are indicated by gene symbols above peaks. QQ plots for each phenotype may be found in Figure VIII in the Data Supplement.

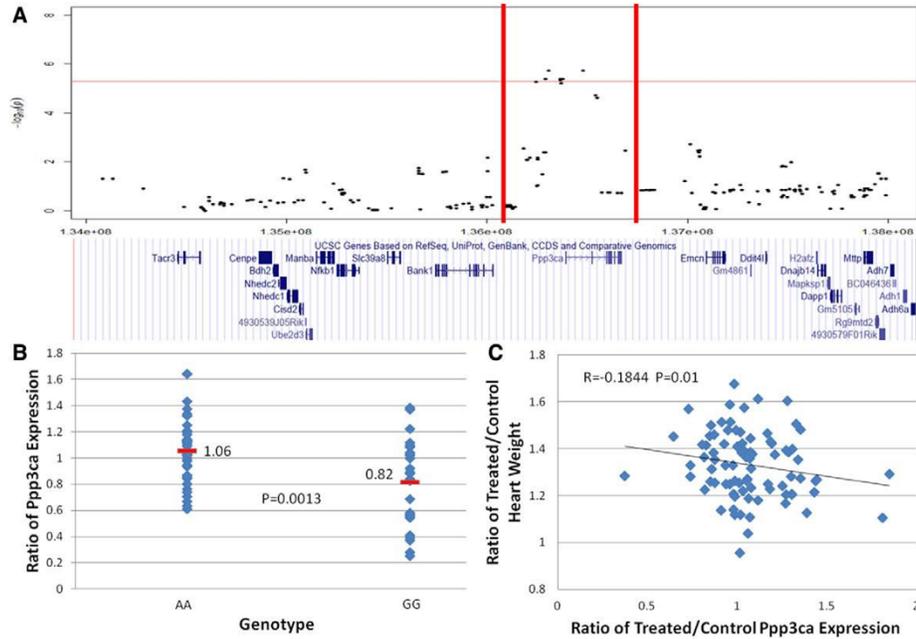
Calcineurin A is the only gene in LD with the peak SNP and has a significant *cis*-eQTLs ( $P=1.3E-3$ ) for the ratio of treated to control calcineurin A expression (Figure 3A and 3B). We also observed a modest correlation between the ratio of *Ppp3ca* expression and the ratio of heart weights in control and ISO-treated animals ( $R=-0.18, P=0.01$ ). We further observed *Ppp3cb*, the beta isozyne of calcineurin A, in a locus on chromosome 14 ( $P=3.3E-6$ ) for the trait treated lung weight. *Ppp3cb* has a strongly suggestive *cis*-eQTL ( $P=4.7E-3$ ) as well as a minor allele with an insertion in a splice site in several strains of the HMDP. In addition to calcineurin A, we identified several other genes with well-established roles in cardiac physiology and pathology within other disease-associated loci. These include the key calcium cycling regulator phospholamban<sup>31</sup> and structural protein *Sgcd*,<sup>32</sup> as well as other genes, which have previously been implicated in cardiac hypertrophy, such as *Prkag2*,<sup>33</sup> or cardiac malformation, such as *Mospd3*<sup>34</sup> (Tables 1 and 2).

The Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>35,36</sup> was used to examine the genes located within LD of the peak SNPs of our study. We observed significant enrichment for calmodulin-related genes

(benjamini-corrected  $P=0.02$ ) and suggestive enrichment categories, such as calcium signaling (uncorrected  $P=0.003$ ) and Epidermal Growth Factor signaling ( $P=0.009$ ), both of which lie downstream of  $\beta$ -adrenergic signaling (Table VI in the Data Supplement). Examination of our top candidate genes within each locus reveals a strong bias toward genes known to be involved in catecholamine-stimulated cardiomyopathy, such as calcineurin, phospholamban, and calmodulin (Tables 1 and 2)

### Conservation of Cardiomyopathy Loci in Mice and Humans

We explored whether the loci we identified overlap with human GWAS results by examining the top 12 previously identified significant and suggestive human loci.<sup>6,7,37</sup> The human loci were mapped onto the mouse genome using the NCBI Homologene resource and compared with a set of loci identified for the weight traits based on a slightly relaxed stringency ( $P<1E-05$ , minor allele frequency  $>5\%$ ) from the HMDP study. We observed 6 out of 12 human loci, including one of the genome-wide significant loci near USP3, replicating in our study (Table IV in the Data Supplement). We determined that this overlap is highly significant ( $P=3.5E-4$ )



**Figure 3.** *Ppp3ca* (calcineurin A) is a candidate gene at the Chromosome 3 right ventricular weight ratio locus. **A**, *Ppp3ca* is the only gene within the locus. The red horizontal line represents the significance threshold, whereas the red vertical lines indicate the limits of the linkage disequilibrium block. The dots represent single-nucleotide polymorphisms, plotted in basepairs along the chromosome with  $-\log_{10}$  ( $P$  value) on the Y-axis. Below are the locations of genes from a genome browser. **B**, *Ppp3ca* has a significant *cis*-expression quantitative trait loci (eQTL;  $P=1.3E-3$ ) for the ratio of gene expression (after and before treatment). Red line and number indicate average gene expression. **C**, The ratio of *Ppp3ca* expression has significant ( $P=0.01$ ) negative correlation with the ratio of treated to untreated heart weights.

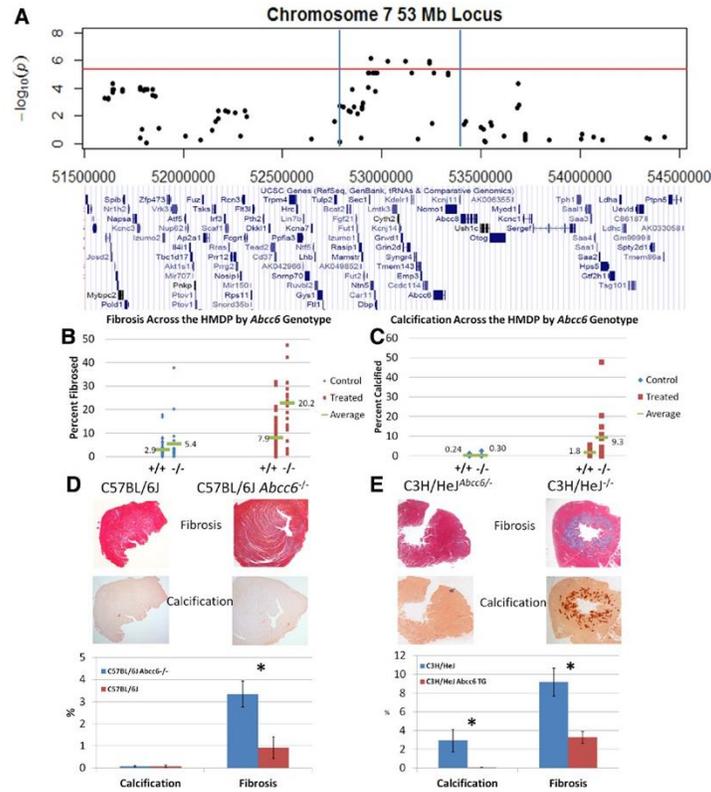
by permutation analysis. Our result supports the concept that genetic influences in  $\beta$ -adrenergic signaling significantly contribute to polygenic human HF. Furthermore, 16 loci for HF-related traits have previously been identified via linkage studies in mice for HF-related traits. We further observed that 7 of these 16 QTLs overlap with weight loci identified in this study ( $P=1.2E-3$  by permutation analysis) (Table V in the Data Supplement), despite the fact that some of the linkage studies used different hypertrophy-inducing stressors, such as a calsequestrin transgene,<sup>8</sup> which likely influence distinct HF pathways.

#### Validation of *Abcc6* as Causal Gene for ISO-Induced Cardiac Fibrosis

We have identified a locus contributing to ISO-induced fibrosis on chromosome 7 (Figure 4,  $P=7.1E-7$ ). One of the 28 genes within the LD block, *Abcc6*, has a splice site variation,<sup>38</sup> resulting in a premature stop codon that is found in 19 of the strains we analyzed for cardiac fibrosis (KK/Hij, C3H/HeJ, DBA/2J, 11 BxD, 2 BxH, 3 CxB) in the HMDD. In untreated animals, we did not observe any significant difference ( $P=0.25$ ) between the degrees of fibrosis present in the left ventricle among HMDD strains divided based on *Abcc6* genotype. In contrast, we observed a marked increase in cardiac fibrosis in the mice containing the *Abcc6* splice mutation allele in response to ISO treatment ( $P=1E-4$ ; Figure 4B).

*Abcc6* deficiency is the cause of pseudoxanthoma elasticum, a disorder characterized by progressive tissue calcification,<sup>13,39</sup> and a deficiency of *Abcc6* has previously been linked to calcification phenotypes in aged mice by our laboratory.<sup>38</sup> However, the fibrosis phenotype observed in these studies is clearly distinct from that of calcification. In fact, we did not observe a significant association with heart calcification in our study under basal ( $P=0.8$ ) or ISO treatment condition ( $P=0.12$ ) in mice with different *Abcc6* genotypes, although we did observe some outlier strains, such as KK/Hij, that had markedly increased calcification after ISO stimulation (Figure 4C). Therefore, *Abcc6* is a likely candidate gene contributing to ISO-induced fibrosis in heart.

To validate the role of *Abcc6* in cardiac fibrosis, we studied a previously described *Abcc6* KO mouse carrying a targeted mutation in a C57BL/6J strain background (KO), as well as the wildtype C57BL/6J (Control) mice.<sup>13</sup> We previously reported that the *Abcc6* KO mice exhibited increased cardiac calcification beginning from 6 months of age.<sup>14</sup> At 3 months of age, neither the wildtype nor the KO mice exhibited significant differences in calcification as judged by Alizarin Red staining or cardiac fibrosis based on Masson Trichrome staining in the absence of ISO treatment (Figure 4D). Consistent with our observations of the entire panel, the ISO treatment significantly increased fibrosis levels in the *Abcc6* KO animals without significantly increasing calcification levels in these



**Figure 4.** *Abcc6* plays a role in the regulation of cardiac fibrosis after isoproterenol (ISO) stimulation. **A**, The locus on chromosome 7 which contains *Abcc6* spans ~800 kb and contains 28 genes within LD. **B**, Calcification in post-ISO treated hearts is increased in mice lacking *Abcc6*. **C**, Expression of *Abcc6* in a mouse that lacks the gene is sufficient to rescue the mouse from the ISO-induced fibrosis and calcification. **D**, Knockout of *Abcc6* in a mouse is sufficient to cause ISO-induced fibrosis, but does not cause a significant increase in calcification. WT, n=4; KO, n=3. **E**, The expression of *Abcc6* as a transgene is sufficient to significantly reduce both calcification and fibrosis. TG, n=8; KO, n=6 (\**P*<0.05).

mice. This result suggests that the age-associated calcification phenotype observed in *Abcc6* KO is distinct from the ISO-induced cardiac fibrosis phenotype.

To further establish *Abcc6* as a causal gene for ISO-induced cardiac fibrosis, we studied transgenic mice carrying the *Abcc6* wildtype locus from a C57BL/6J bacterial artificial chromosome on the background of the fibrosis-susceptible C3H/HeJ strain. Strain C3H/HeJ mice lack functional *Abcc6* as a result of a splice variation.<sup>40</sup> In the absence of ISO, neither C3H/HeJ mice nor C3H/HeJ mice carrying the *Abcc6* bacterial artificial chromosome-transgene exhibited significant calcification or fibrosis in the heart, whereas after ISO treatment, the C3H/HeJ mice but not the *Abcc6*-bacterial artificial chromosome transgenic mice exhibited substantial fibrosis and calcification (Figure 4E). The differing results between the C57BL/6J KO mice and the naturally mutant C3H/HeJ mice suggests that additional modifier genes are necessary to induce cardiac calcification in *Abcc6* KO mice after ISO stimulation, whereas *Abcc6* KO is sufficient to cause ISO-induced cardiac fibrosis. The mechanisms by which *Abcc6* promotes fibrosis

is unknown, but DAVID analysis of genes significantly correlated with *Abcc6* expression in heart showed highly significant enrichment for mitochondrial genes (Table VII in the Data Supplement). Systemic factors are clearly involved in the calcification phenotype of *Abcc6* deficiency,<sup>40</sup> but it is noteworthy that, based on our expression profiling data, *Abcc6* is clearly expressed in heart. The *Abcc6* locus is suggestively associated (*P*<0.01) with several hypertrophic phenotypes in our study, namely whole heart weight (*P*=4E-3), right atrial weight (*P*=5E-3), and LVW (*P*=7E-3). Transcript coexpression network analysis was performed on *Abcc6*, but the module into which *Abcc6* fell was not significantly enriched for any particular Gene Ontology term using DAVID (data not shown).

### Discussion

We have used a strategy involving GWAS in a large panel of inbred mice to perform fine mapping of loci contributing to specific pathological features of cardiomyopathy after treatment with ISO.<sup>26,27</sup> We have combined this strategy with global gene expression analysis in the heart to help identify

candidate genes. A significant number of genetic loci revealed from this study are replicated in human GWAS analysis, supporting a conserved genetic network contributing to human heart failure. Several loci contain genes known to be involved in cardiomyopathy based on previous biochemical or genetic studies, supporting the validity of this approach to uncover important mechanisms and pathways related to the onset of heart failure. Finally, we validated *Abcc6*, a candidate GWAS hit, as a novel player in stress-induced cardiac fibrosis. These findings should complement human studies to identify genes and pathways contributing to this common and poorly understood disorder.

Systems genetics is a potent approach to reveal genes and pathways underlying the specific pathological features of cardiomyopathies, of which GWAS represents only one potential avenue for exploration. Using the resources presented here, it should be possible to perform additional analyses, including the generation of transcript coexpression networks for the identification of groups of genes involved in maladaptive cardiac remodeling. The genetic information and the phenotypic spectra established by this study should provide a valuable resource for future heart failure studies.

### Acknowledgments

We thank the excellent technical assistance from Ms. Mary Tuteryan, Ms. Haiying Pu, and Ms. Melenie Rosales.

### Sources of Funding

This work was supported by National Institute of Health (NIH) grants HL110667 and HL28481. C.D. Rau was supported by NIH training grant T32HL69766 and J. Wang was supported by NIH training grant HL007895.

### Disclosures

None.

### References

- Mudd JO, Kass DA. Tackling heart failure in the twenty-first century. *Nature*. 2008;451:919–928. doi: 10.1038/nature06798.
- Beltrami CA, Finato N, Rocco M, Feruglio GA, Puricelli C, Cigola E, et al. Structural basis of end-stage failure in ischemic cardiomyopathy in humans. *Circulation*. 1994;89:151–163.
- Frangogiannis NG. Regulation of the inflammatory response in cardiac repair. *Circ Res*. 2012;110:159–173. doi: 10.1161/CIRCRESAHA.111.243162.
- Kong P, Christia P, Frangogiannis NG. The pathogenesis of cardiac fibrosis. *Cell Mol Life Sci*. 2014;71:549–574. doi: 10.1007/s00018-013-1349-6.
- Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J*. 2011;32:1065–1076.
- Morrison AC, Felix JF, Cupples LA, Glazer NL, Loehr LR, Dehghan A, et al. Genomic variation associated with mortality among adults of European and African ancestry with heart failure: the cohorts for heart and aging research in genomic epidemiology consortium. *Circ Cardiovasc Genet*. 2010;3:248–255. doi: 10.1161/CIRCGENETICS.109.895995.
- Parsa A, Chang YP, Kelly RJ, Corretti MC, Ryan KA, Robinson SW, et al. Hypertrophy-associated polymorphisms ascertained in a founder cohort applied to heart failure risk and mortality. *Clin Transl Sci*. 2011;4:17–23. doi: 10.1111/j.1752-8062.2010.00251.x.
- Le Corvoisier P, Park HY, Carlson KM, Marchuk DA, Rockman HA. Multiple quantitative trait loci modify the heart failure phenotype in murine cardiomyopathy. *Hum Mol Genet*. 2003;12:3097–3107. doi: 10.1093/hmg/ddg333.
- McDermott-Roe C, Ye J, Ahmed R, Sun XM, Serafin A, Ware J, et al. Endonuclease G is a novel determinant of cardiac hypertrophy and mitochondrial function. *Nature*. 2011;478:114–118. doi: 10.1038/nature10490.
- Wheeler FC, Fernandez L, Carlson KM, Wolf MJ, Rockman HA, Marchuk DA. QTL mapping in a mouse model of cardiomyopathy reveals an ancestral modifier allele affecting heart function and survival. *Mamm Genome*. 2005;16:414–423. doi: 10.1007/s00335-005-2468-7.
- Wheeler FC, Tang H, Marks OA, Hadnoti TN, Chu PL, Mao L, et al. Tnfr3 modifies disease progression in murine models of cardiomyopathy. *PLoS Genet*. 2009;5:e1000647. doi: 10.1371/journal.pgen.1000647.
- Bennett BJ, Farber CR, Orozco L, Kang HM, Ghazalpour A, Siemers N, et al. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res*. 2010;20:281–290. doi: 10.1101/gr.099234.109.
- Mungruc IN, Zhao P, Yao Y, Meng H, Rau C, Havel JV, et al. *Abcc6* deficiency causes increased infarct size and apoptosis in a mouse cardiac ischemia-reperfusion model. *Arterioscler Thromb Vasc Biol*. 2011;31:2806–2812. doi: 10.1161/ATVBAHA.111.237420.
- Martin LJ, Lau E, Singh H, Vergnes L, Tarling EJ, Mehrabian M, et al. *ABCC6* localizes to the mitochondria-associated membrane. *Circ Res*. 2012;111:516–520. doi: 10.1161/CIRCRESAHA.112.276667.
- Berk BC, Fujiwara K, Lehoux S. ECM remodeling in hypertensive heart disease. *J Clin Invest*. 2007;117:568–575. doi: 10.1172/JCI131044.
- Bronson RT. Rate of occurrence of lesions in 20 inbred and hybrid genotypes of rats and mice sacrificed at 6-month intervals during the first years of life. In: Harrison D, ed. *Genetic Effects of Aging II*. Caldwell, NJ: Tel-ford Press; 1990:279–358.
- Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient control of population structure in model organism association mapping. *Genetics*. 2008;178:1709–1723. doi: 10.1534/genetics.107.080101.
- Smyth GK. No Title. In: Gentleman R, Irizarry W, Huber W, ed. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York: Springer; 2005:397–420.
- Johnson WF, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8:118–127. doi: 10.1093/biostatistics/kxj037.
- Parks BW, Nam E, Org E, Kostem E, Norheim F, Hui ST, et al. Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab*. 2013;17:141–152. doi: 10.1016/j.cmet.2012.12.007.
- Grundberg E, Small KS, Hedman ÅK, Nica AC, Buil A, Keildson S, et al.; Multiple Tissue Human Expression Resource (MuTHER) Consortium. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet*. 2012;44:1084–1089. doi: 10.1038/ng.2394.
- Borel C, Deutsch S, Letourneau A, Migliavacca E, Montgomery SB, Dimas AS, et al. Identification of cis- and trans-regulatory variation modulating microRNA expression levels in human fibroblasts. *Genome Res*. 2011;21:68–73. doi: 10.1101/gr.109371.110.
- Nica AC, Ongen H, Irminger JC, Bosco D, Berny T, Antonarakis SF, et al. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res*. 2013;23:1554–1562. doi: 10.1101/gr.150706.112.
- Hasin-Brumshtein Y, Hormozdiari F, Martin L, van Nas A, Eskin E, Lusis AJ, et al. Allele-specific expression and eQTL analysis in mouse adipose tissue. *BMC Genomics*. 2014;15:471. doi: 10.1186/1471-2164-15-471.
- Dorn GW III, Liggett SB. Pharmacogenomics of beta-adrenergic receptors and their accessory signaling proteins in heart failure. *Clin Transl Sci*. 2008;1:255–262. doi: 10.1111/j.1752-8062.2008.00059.x.
- Breckenridge R. Heart failure and mouse models. *Dis Model Mech*. 2010;3:138–143. doi: 10.1242/dmm.005017.
- Zhang X, Szeto C, Gao E, Tang M, Jin J, Fu Q, et al. Cardiotoxic and cardioprotective features of chronic  $\beta$ -adrenergic signaling. *Circ Res*. 2013;112:498–509. doi: 10.1161/CIRCRESAHA.112.273896.
- Berthoumce C, Peter B, Schüpfer F, Hayoz P, Kutalik Z, Abriel H, et al. Cardiovascular response to beta-adrenergic blockade or activation in 23 inbred mouse strains. *PLoS One*. 2009;4:e6610. doi: 10.1371/journal.pone.0006610.
- Flint J, Eskin E. Genome-wide association studies in mice. *Nat Rev Genet*. 2012;13:807–817. doi: 10.1038/nrg3335.
- Yalcin B, Wong K, Agam A, Goodson M, Keane TM, Gan X, et al. Sequence-based characterization of structural variation in the mouse genome. *Nature*. 2011;477:326–329. doi: 10.1038/nature10432.
- Schönberger J, Seidman CE. Many roads lead to a broken heart: the genetics of dilated cardiomyopathy. *Am J Hum Genet*. 2001;69:249–260. doi: 10.1086/321978.

Downloaded from <http://circgenetics.ahajournals.org/> at CONS CALIFORNIA DIG LIB on May 9, 2016

32. Bauer R, Macgowan GA, Blain A, Bushby K, Straub V. Steroid treatment causes deterioration of myocardial function in the  $\delta$ -sarcoglycan-deficient mouse model for dilated cardiomyopathy. *Cardiovasc Res*. 2008;79:652–661. doi: 10.1093/cvr/cvn131.
33. Banerjee SK, McGaffin KR, Huang XN, Ahmad F. Activation of cardiac hypertrophic signaling pathways in a transgenic mouse with the human PRKAG2 Thr400Asn mutation. *Biochim Biophys Acta*. 2010;1802:284–291. doi: 10.1016/j.bbdis.2009.12.001.
34. Pall GS, Wallis J, Axton R, Brownstein DG, Gautier P, Buerger K, et al. A novel transmembrane MSP-containing protein that plays a role in right ventricle development. *Genomics*. 2004;84:1051–1059. doi: 10.1016/j.ygeno.2004.08.017.
35. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003;4:P3.
36. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44–57. doi: 10.1038/nprot.2008.211.
37. Ellinor PT, Sasse-Klaassen S, Probst S, Gerull B, Shin JT, Toepffel A, et al. A novel locus for dilated cardiomyopathy, diffuse myocardial fibrosis, and sudden death on chromosome 10q25-26. *J Am Coll Cardiol*. 2006;48:106–111. doi: 10.1016/j.jacc.2006.01.079.
38. Meng H, Vera I, Cho N, Wang X, Wang SS, Ingram-Drake L, et al. Identification of *Abcc6* as the major causal gene for dystrophic cardiac calcification in mice through integrative genomics. *Proc Natl Acad Sci U S A*. 2007;104:4530–4535. doi: 10.1073/pnas.0607620104.
39. Campens L, Vanakker OM, Traehet B, Segers P, Leroy BP, De Zaeytjijd J, et al. Characterization of cardiovascular involvement in pseudoxanthoma elasticum families. *Arterioscler Thromb Vasc Biol*. 2013;33:2646–2652. doi: 10.1161/ATVBAHA.113.301901.
40. Jiang Q, Oldenburg R, Otsuru S, Grand-Pierre AE, Horwitz EM, Uitto J. Parabolic heterogenetic pairing of *Abcc6*<sup>-/-</sup>/*Rag1*<sup>-/-</sup> mice and their wild-type counterparts halts ectopic mineralization in a murine model of pseudoxanthoma elasticum. *Am J Pathol*. 2010;176:1855–1862. doi: 10.2353/ajpath.2010.090983.

### CLINICAL PERSPECTIVE

Heart failure (HF) is characterized by insufficient pump action to maintain adequate blood flow, and common forms can result from myocardial infarction, hypertension, valvular disease, and viral infection. It constitutes the major cause of hospitalization in the United States with a lifetime risk of  $\approx 1$  in 5. Genetic approaches to understand the common forms of HF, including genome-wide association studies have been largely unsuccessful, likely because of the extremely heterogeneous pathogenesis of HF. Using a defined pathological stress, chronic stimulation by isoproterenol, to circumvent the problem of heterogeneity, we performed genome-wide association studies for genes influencing cardiac hypertrophy and fibrosis in a large panel of inbred mice. We identified 24 significant or suggestive loci, several of which contain known HF-related genes, such as phospholamban and calcineurin A. We also validated *Abcc6* as a causal gene at one of the fibrosis loci. Our results have the potential to help identify novel pathways leading to human HF and to develop novel therapies.

## **Chapter 4**

### **Genetic Dissection of Cardiac Remodeling in an Isoproterenol-induced Heart Failure Mouse Model.**

**Chapter 4 Preface:** This chapter is a manuscript in press at PLOS Genetics that describes a population based genetic screen using inbred mouse strains (Hybrid Mouse Diversity Panel – HMDP) to identify novel regulators of heart failure phenotypes. In this project our group identified the gene *Myh14* as a contributor to left ventricular dilation. As a collaborator on this project I used siRNA knockdown to alter the expression of *Myh14* in neonatal rat ventricular cardiomyocytes to determine if changes in *Myh14* expression altered cardiomyocyte size or viability in response to adrenergic stimulus. In addition to these *in vitro* experiments, I performed correlation and eQTL analysis to identify downstream targets of *Myh14* in the heart. Using this combination approach I identified the gene *Foxo1* as a putative downstream target of *Myh14* that was then validated in-vivo using *Myh14* knock out mice.

## Title

Genetic Dissection of Cardiac Remodeling in an Isoproterenol-induced Heart Failure Mouse Model: A genome-wide association study of heart failure in mice

## Authors

Jessica Jen-Chu Wang<sup>1\*</sup>, Christoph Rau<sup>2</sup>, Rozeta Avetisyan<sup>2</sup>, Shuxun Ren<sup>2</sup>, Milagros Romay<sup>3</sup>, Gabriel Stolin<sup>4</sup>, Kei Wei Gong<sup>1</sup>, Yibin Wang<sup>1,2,5¶</sup>, Aldons J. Lusis<sup>1,3,6¶\*</sup>

## Affiliations

<sup>1</sup>Division of Cardiology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, United States of America

<sup>2</sup>Department of Anesthesiology, David Geffen School of Medicine, University of California, Los Angeles, United States of America

<sup>3</sup>Department of Microbiology, Immunology, and Molecular Genetics, David Geffen School of Medicine, University of California, Los Angeles, United States of America

<sup>4</sup>Department of Molecular, Cell, and Developmental Biology, David Geffen School of Medicine, University of California, Los Angeles, United States of America

<sup>5</sup>Department of Physiology, David Geffen School of Medicine, University of California, Los Angeles, United States of America

<sup>6</sup>Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, United States of America

\* Corresponding authors:

E-mail: [jessicawang@mednet.ucla.edu](mailto:jessicawang@mednet.ucla.edu) (JJW)

E-mail: [jlusis@mednet.ucla.edu](mailto:jlusis@mednet.ucla.edu) (AJL)

¶YW and AJL are Joint Senior Authors.

## Abstract

We aimed to understand the genetic control of cardiac remodeling using an isoproterenol-induced heart failure model in mice, which allowed control of confounding factors in an experimental setting. We characterized the changes in cardiac structure and function in response to chronic isoproterenol infusion using echocardiography in a panel of 104 inbred mouse strains. We showed that cardiac structure and function, whether under normal or stress conditions, has a strong genetic component, with heritability estimates of left ventricular mass between 61% and 81%. Association analyses of cardiac remodeling traits, corrected for population structure, body size and heart rate, revealed 17 genome-wide significant loci, including several loci containing previously implicated genes. Cardiac tissue gene expression profiling, expression quantitative trait loci, expression-phenotype correlation, and coding sequence variation analyses were performed to prioritize candidate genes and to generate hypotheses for downstream mechanistic studies. Using this approach, we have validated a novel gene, *Myh14*, as a negative regulator of ISO-induced left ventricular mass hypertrophy in an *in vivo* mouse model and demonstrated the up-regulation of immediate early gene *Myc*, fetal gene *Nppb*, and fibrosis gene *Lgals3* in ISO-treated *Myh14* deficient hearts compared to controls.

## **Introduction**

Heart failure (HF) is a major health issue, affecting 5.7 million people in the United States (1). Despite advances in therapy, HF remains a lethal condition with 5- and 10-year mortality rates of greater than 40% and 60% (2). A number of etiologic factors, such as coronary artery disease, hypertension, valvular disease, alcohol, chemotherapy, and rare deleterious genetic mutations can lead to cardiac injury that results in HF but little is known about how common genetic variants contribute to HF progression. Irrespective of the primary insult, compensatory adrenergic and renin-angiotensin activation augment heart rate (HR), contractility and fluid retention to maintain adequate cardiac output and preserve organ function, which in turn leads to chronic maladaptive cellular growth and irreversible myocardial injury, furthering HF progression (3).

Such molecular, cellular and extracellular changes, manifested clinically as changes in size, shape and function of the heart, is also known as cardiac remodeling and is one of the most important clinical determinants of HF progression. In addition,  $\beta$ -adrenergic receptor blockers and angiotensin-converting enzyme (ACE) inhibitors, HF therapeutic agents that provide morbidity and mortality benefits, reverse ventricular dilation and systolic dysfunction, further supporting the importance of clinical cardiac remodeling as HF therapeutic targets (4, 5). Understanding how common genetic variation modifies the pathophysiology of HF progression in terms of cardiac remodeling will likely provide insights in the design of novel therapeutics to improve survival and life quality of HF patients.

The unbiased genome-wide association study (GWAS) design is well suited to detect the effects of genetic variation on complex traits such as HF (6). However, a number of human HF GWAS performed to-date have had limited success. For example, a study on cardiac structure and function yielded one significant locus explaining <1% of variance in left ventricular diastolic diameter (LVIDd) (7). A meta-analysis of 4 community-based GWAS cohorts, involving nearly 24,000 subjects, identified only two loci to be significantly associated with incident HF, explaining a very small fraction of the variance (8). A sporadic dilated cardiomyopathy GWAS, involving 1179 cases and 1108 controls from several European populations, identified only two associated loci (9). Lastly, a GWAS of cardiac structure and function in 6,765 African Americans identified 4 loci associated with left ventricular mass (LVM), interventricular septal wall thickness (IVSd), LVIDd, and ejection fraction (EF) (10). In spite of their scales, the paucity of detailed phenotypic data as well as environmental heterogeneity hampered detection of genetic signals driving HF in these large human cohorts.

The laboratory mouse, with its fully sequenced and annotated genome, targeted germ line modification and easy accessibility, provides a means of overcoming some of the challenges in human studies and is complementary. Particularly informative have been reverse genetic studies of candidate genes and pathways utilizing engineered mouse models. Studies of natural variation of mice and rats, some involving sensitized models, have also resulted in the identification of novel pathways contributing to HF and other common cardiovascular traits (11-15). A major difficulty of the latter, however, has been the poor resolution of classical linkage studies, making the identification of underlying causal genes a laborious process.

Recently, with the development of high-density genotyping and sequencing in rodents, relatively high-resolution association mapping approaches, analogous to human GWAS, have

become feasible (16). Our group has pioneered a resource termed the Hybrid Mouse Diversity Panel (HMDP), a panel of 100+ strains of inbred mice that have either been sequenced or densely genotyped and display natural inter-strain genetic variation, allowing a mapping resolution more than an order of magnitude higher than traditional crosses (17). The method combines the use of classic inbred strains for mapping resolution and recombinant inbred (RI) strains for power and has been used to successfully identify a number of genes and loci involved in lipid, obesity, bone, and behavioral traits (16, 18-20). Because the HMDP strains are renewable, the resource is well suited to the application of systems genetics approaches, involving the integration of high throughput molecular phenotypes, such as expression array data, with clinical phenotypes.

We have previously reported a genetic analysis of  $\beta$ -adrenergic agonist isoproterenol (ISO)-induced cardiac hypertrophy in the HMDP, identifying a number of loci contributing to heart weight and fibrosis (21). We now extend this study to examine ISO-induced cardiac remodeling, i.e. cardiac structural and functional changes as characterized by serial echocardiography, a powerful and noninvasive tool to serially monitor cardiac structure and function in murine injury models (22, 23). We demonstrate fine phenotypic data across the concentric-eccentric and functional cardiac remodeling spectra and found that cardiac remodeling traits among the HMDP are highly heritable. Using a common genetic variation association method, we identify 17 genome-wide significant loci with high resolution that modify ISO-induced cardiac remodeling in the heart, many of which are less than 1 Mb in size. Then, we utilize expression arrays from left ventricular (LV) tissues, with and without ISO treatment, to understand the genetic control of gene expression and gene expression correlation to cardiac remodeling phenotypes to further prioritize candidate genes in each locus. Finally,

using *in vitro* and *in vivo* models, we validate a novel gene *Myh14* as causative in ISO-induced cardiac remodeling not previously reported in our heart weight analyses (21).

## Materials and Methods

### **Chronic $\beta$ -adrenergic stimulation, echocardiography, and tissue collection from the HMDP**

**HMDP mice.** Breeding pairs from the HMDP inbred strains were obtained from the Jackson Laboratory (Bar Harbor, ME, USA). Eight- to ten-week-old nulliparous female offspring from the following 104 mouse strains were used, including 30 classical inbred strains (129X1/SvJ, A/J, AKR/J, BALB/cByJ, BALB/cJ, BTBRT<+>tf/J, BUB/BnJ, C3H/HeJ, C57BL/6J, C57BLKS/J, C57L/J, C58/J, CBA/J, CE/J, DBA/2J, FVB/NJ, KK/HIJ, LG/J, LP/J, MA/MyJ, NOD/LtJ, NON/LtJ, NZB/BINJ, NZW/LacJ, PL/J, RIIIS/J, SEA/GnJ, SJL/J, SM/J, SWR/J) and 74 recombinant inbred (RI) strains [RI (number of strains) – AXB (8), BXA (10), BXD (44), BXH(5), CXB (7)]. All animal experiments were conducted following guidelines established and approved by the University of California, Los Angeles Institutional Animal Care and Use Committee.

**Chronic  $\beta$ -adrenergic stimulation and tissue collection.** Using i.p. ketamine as a surgical anesthetic agent, ALZET Model 1004 minipumps (Cupertino, CA, USA) were implanted intraperitoneally to administered ISO, a non-selective  $\beta$ -adrenergic agonist that has been used widely to mimic the HF state in experimental animals (24, 25), at a dose of 30 mg/kg body weight/day for 21 days. Approximately four ISO-treated (S5 Table) and 2-4 control ((21) Table S1) mice from each strain were examined. At the end of the protocol, mice were sacrificed by giving a sub-lethal dosage of inhaled isoflurane followed by cervical dislocation. LV tissues were collected and frozen immediately in liquid nitrogen.

Echocardiography. To ensure adequate sedation while minimizing the effects of inhaled isoflurane on loading conditions, HR, cardiac structure and function, we minimized induction and maintenance doses of isoflurane at or below 1.25% and 1%, respectively, while closely

monitoring for HR < 475 bpm as a sign for deep sedation and adjusting isoflurane dosage as needed (26).

A single operator (JJW), who followed a standard operating protocol detailed below, performed all of the echocardiograms using the Vevo 770 ultrasound system (VisualSonics, Inc., Toronto, ON, Canada): A parasternal long-axis B-mode image was obtained. The maximal long-axis of the LV was positioned perpendicular to the ultrasound beam. A 90° rotation of the ultrasound probe at the papillary muscle level was performed to obtain a parasternal short-axis view of the LV. A M-mode image to document LV dimensions was captured and saved for analysis. At a later time, saved images were analyzed using the Vevo 770 cardiac analysis package by a single observer (JJW) who was blinded to mouse strains.

A baseline echocardiogram was performed on all the mice. In ISO-treated mice, serial echocardiograms were performed at 1, 2, and 3 weeks of treatment. In control mice, a second echocardiogram was performed in 70 mouse strains at week 3 as an internal control. The reproducibility of our echocardiographic measurements was assessed using the Bland-Altman plots, which demonstrated acceptable agreement between week 0 and week 3 control measurements (S1B Fig). IVSd and LVM were not statistically different between the 2 time points, although fractional shortening (FS) was 2.3% higher and LVIDd 0.07 mm lower at week 3 compared to baseline ( $p=0.002$  and  $0.019$ , respectively), possibly due to the effects of co-housing with ISO-treated animals (S1 Table).

### **Transcriptome profiling and analysis**

**RNA extraction.** Frozen LV tissues were homogenized in QIAzol Lysis Reagent prior to RNA isolation using RNeasy columns (QIAGEN, Valencia, CA, USA). RNA quality was assessed

using the Bioanalyzer RNA kits (Agilent Technologies, Santa Clara, CA, USA). RIN  $\geq 7.0$  was considered acceptable.

**Expression array profiling.** Expression profiling of pooled biological replicates was performed using Illumina MouseRef-8 v2.0 Expression BeadChip arrays (Illumina, Inc., San Diego, CA, USA). Raw data were deposited in the Gene Expression Omnibus (GEO) online database <http://www.ncbi.nlm.nih.gov/geo/> (Accession GSE48760). To minimize artifacts due to single nucleotide polymorphisms (SNPs), we excluded probes that aligned to sequences containing known SNPs. Background correction and quantile normalization of the image data was performed using the `neqc` method from the R package `limma` (27). Hierarchical clustering of samples was performed to exclude outlier samples using the R package `WGCNA` (28). In total, expression profiles for 90 control and 91 ISO strains (including 82 strains with matching control and isoproterenol samples) were included in downstream analyses.

**Differential gene expression.** Using the R package `limma`, moderated t-statistics and the associated p values were calculated. Multiple testing was corrected by controlling for false discovery rate using the Benjamini-Hochberg procedure (29). Probes with log<sub>2</sub>-fold change  $> 0.2$  and adjusted p-value  $< 0.05$  were considered significantly differentially expressed. Functional analysis of probe lists was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (30) to identify pathways and cellular processes enriched in genes differentially regulated by ISO stimulation (30, 31). In total 1,502 of the 18,335 probes (8.2%) were differentially expressed at  $> 15\%$  change at an adjusted p-value  $< 0.05$ , including 840 up-regulated and 662 down-regulated probes (S2A Table). DAVID gene ontology analysis was performed on the differentially expressed probes as well as the top 1000 correlated transcripts for each phenotype-expression pair (S2B-D Tables).

## **Statistical analyses, heritability estimation, association mapping and candidate prioritization**

**Statistical analysis and graphical display.** The standard R program was used to performed t-test. The R package vioplot was used to plot violin plots. The R package corrplot was used to plot correlations among traits. Consistent with time series data, serial measurements of the same trait were generally correlated to each other with adjacent time points being the most correlated. Of note, LVM was a calculated measure based in part on IVSd and LVIDd; therefore, it is not surprising that LVM was correlated with IVSd and LVIDd at each time point. The R package FactoMineR was used to compute principal component analysis and display variables factor map and individuals factor map. The R package WGCNA was used to calculate biweight midcorrelation (bicor), which is a median based correlation measure that is more robust than the Pearson correlation but often more powerful than the Spearman correlation. The proportion of variance ( $\omega^2$ ) explained was computed from the formula:

$$\omega^2 = \frac{SSQ_{condition} - (k - 1)MSE}{SSQ_{total} + MSE} \quad (1)$$

where MSE is the mean square error and k is the number of conditions.

Genotypes. The HMDP mouse strains were previously genotyped using the JAX Mouse Diveristy Genotyping Array (32, 33). To select for informative and high quality SNPs, each SNP was filtered for > 5% minor allele frequency and < 10% missing values among the strains using plink (34). Approximately 210,000 of the informative SNPs were selected.

**Heritability calculation.** Heritability is the proportion of observed differences due to genetic variation in the population. Specifically, broad-sense heritability ( $H^2$ ) reflects all the genetic contributions to a population's phenotypic variance including additive, dominant, and epistatic effects, while narrow-sense heritability ( $h^2$ ) represents the proportion of phenotypic variance that is due to additive genetic effects alone. In inbred model organisms, repeatability expresses the proportion of variation in a trait that is due to differences between individual strains and not due to differences within individual strains. Under the assumption that all differences between genotypes are genetic and not due to genotype-environment correlation, repeatability is equal to broad-sense heritability; otherwise it only provides an upper-bound for broad-sense heritability. Using the R package heritability (35), an analysis of variance (ANOVA)–based method that accounts for differing numbers of replicates, called line repeatability, was used to estimate  $H^2$  for each phenotype and time point. In addition, a linear model method incorporating genetic relatedness matrix and within-strain variability, called marker-based  $h^2$ , was used to compute  $h^2$  and confidence intervals.

**Genome-wide significance threshold.** The threshold for genome-wide significance was determined by simulation and permutation of strain genotypes, as previously described (19). The significance threshold for clinical traits was determined to be  $4.1 \times 10^{-6}$ . The significance threshold for cis-eQTL and trans-eQTL were  $3.63 \times 10^{-3}$  and  $6.1 \times 10^{-5}$ , respectively. Given that the studied traits could be collapsed to roughly a handful of PCs, the significance threshold for clinical trait analyses was kept at the nominal p-value of  $4.1 \times 10^{-6}$ , without Bonferroni adjustment for multiple testing.

**Association mapping.** We used the Factored Spectrally Transformed Linear Mixed Models (FaST-LMM) to test for association while accounting for population structure and genetic relatedness among strains as previously described (36), which produced the same results as efficient mixed-model association (EMMA) but with a run time and memory footprint that is only linear to the cohort size (S2 Fig). Association mapping was performed for echocardiographic measurements at each time point and the change in measurements from baseline at each time point while adjusting for baseline body weight (BBW) and HR to define clinical quantitative trait loci (cQTL). In addition, association mapping for gene expression traits, under baseline and isoproterenol-treated conditions, was performed to define expression quantitative trait loci (eQTL). As a general rule, upwards of 90 to 100 strains are required to ensure stable mapping results.

**Linkage disequilibrium block boundaries.** Linkage disequilibrium (LD) is the non-random association of alleles generally near one another at a locus. SNPs in strong LD tend to have association p-values that are similar in strength. Genomic boundaries around lead association SNPs were chosen based on flanking SNPs with p-values  $< 1 \times 10^{-5}$  that were no more than 2 Mb apart between nearest consecutive pairs. Lead SNPs without any neighboring SNPs at p-value  $< 1 \times 10^{-5}$  were excluded. LD between lead SNPs and flanking SNPs were calculated and visualized by plotting regional plots using a custom-built of LocusZoom (37) based on our HMDP mice.

**Candidate gene prioritization.** Genes within LD ( $r^2 > 0.8$ ) of the lead SNPs were examined for coding sequence and splice region variations using the Wellcome Trust Mouse Genomes Project (MGP) sequencing database (38) using a vcftools-based pipeline. SIFT score, a functional annotation score for non-synonymous variants, was noted whenever available (39). The

expression profiles of genes within LD or nearby the lead SNPs were further examined for the presence of cis-expression quantitative trait loci (cis-eQTL). When a transcript's cis-eQTL coincides with the clinical trait locus and is correlated with the trait, a causal relationship between the locus, the transcript and the trait may be inferred.

### ***Myh14* functional validation in NRVM and in knock-out mice**

**siRNA knock-down neonatal rat ventricular myocytes.** Neonatal rat ventricular myocytes (NRVMs) were isolated and cultured at the UCLA NRVM core facility using 2-4 day old rats. Myocytes and fibroblasts were separated using Percoll density gradient. NRVMs were plated at 375,000 cells/well in DMEM + 10% FBS + 100 U/mL penicillin + 100 µg/mL streptomycin (p/s) on Day 1. On the morning of Day 2, plated NRVMs were transfected with 30 pmol per well of siRNA (siMyh14#1 - #RNC.RNAI.N001100690.12.1, siMyh14#2 - RNC.RNAI.N001100690.12.6., or scrambled control purchased from IDT) using Lipofectamine 2000 (Life Technologies) and media changed to DMEM + 1% ITS (Insulin, Transferrin and Selenium) without p/s for *Myh14* knockdown experiments. On Day 3, at 24 hours following transfection, cells were treated with 20 µM ISO or 50 µM phenylephrine (PE) for an additional 24 hours prior to bright field imaging on day 4.

MTT Assay was performed using the Vybrant® MTT Cell Proliferation Assay Kit (Life Technologies V13154). Briefly, NRVMs were plated in 96 well plates at a density of 10,000 cells per well and were transfected with either scrambled control or *Myh14* siRNA using Lipofectamine RNAimax overnight. Cells were then treated the next day with 20 µM ISO or 50 µM PE for 24 hours and then viability was assayed using MTT.

**Myh14 knock-out mice.** *Myh14*<sup>-/-</sup> mice generated on a 129/Sv background followed by backcrossing to strain C57BL/6 for more than 7 generations were obtained from the Adelstein lab (40) and backcrossed (*Myh14*<sup>-/-</sup> male x C57BL/6J female) to generate *Myh14*<sup>+/-</sup> mice. *Myh14*<sup>+/-</sup> heterozygous crosses generated wild-type (WT), heterozygous (HET) and homozygous (KO) littermates at normal mendelian ratios (Chi-Square test p = 0.13, S3 Table). Six WT, nine HET, and seven KO twelve-week-old nulliparous female mice were subjected to chronic  $\beta$ -adrenergic stimulation and echocardiogram in the same manner as in the HMDP described above.

**Histology and immunofluorescence microscopy.** Hearts were fixed in 10% paraformaldehyde in PBS. Masson's trichrome staining and immunofluorescence staining were performed as previously described (40). The antibodies used for immunofluorescence staining were N-cadherin (mouse, 1:2000, Invitrogen),  $\beta$ -catenin (mouse, 1:1000, Zymed), and wheat germ agglutinin (WGA). Paraffin embedded heart sections were first dewaxed and rehydrated by standard procedures. Following antigen retrieval in citric acid buffer (10 mM, pH 6.0), the sections were blocked with 1% BSA/5% goat serum in PBS for 1 hour at room temperature and then incubated with Alexa Fluor® 594-WGA (10  $\mu$ g/ml, Invitrogen) and DAPI in blocking solution for 1 hour at room temperature. The slides were then washed and mounted using Prolong Anti-fade mount media (Invitrogen).

## **Results**

### **Experimental design and quality control**

We characterized 104 classical and RI mouse strains by obtaining echocardiographic measurements under the baseline condition and in response to chronic ISO administration for 3 weeks (Figure 1). Global gene expression profiling of LV tissues from control and ISO-treated mice was performed to identify genes whose expression was correlated to HF traits and to identify eQTLs to prioritize candidate genes. We previously reported association mapping of heart weight and fibrosis (21); the present study focuses on echocardiographic measures of clinical significance, including IVSd, LVIDd, LVM and FS, to enable mapping of fine cardiac remodeling phenotypes due to ISO-treatment.

To minimize inter-operator and inter-observer variation, all echocardiograms were performed by a single operator and interpreted by the same observer who was blinded to strain name and treatment assignment. Inhaled isoflurane was titrated to achieve adequate sedation, while maintaining target HR above 475 bpm. The Bland-Altman plot, which shows the agreement between two quantitative methods, demonstrated acceptable differences between ISO-treated LVM (a calculated value based on echo parameters) and LV weight (a weighed value after LV was dissected from the atria and RV) estimates (S3A-C Figs). LVM by echocardiogram was significantly correlated to LV weight at sacrifice ( $r = 0.77$  in control and  $r = 0.85$  in ISO both with  $p\text{-value} < 2.2 \times 10^{-16}$ ), validating our echocardiographic measures externally.

**The HMDP exhibits significant variation in cardiac structure and function both under baseline and ISO-treated conditions**

We observed striking variation in cardiac structure and function among the HMDP strains both under baseline condition and in response to chronic ISO administration (Figure 2 and S4). Importantly, the phenotypic spectrum among the BXD RI lines exceeded that of their parental strains C57BL/6 and DBA/2, consistent with cardiac remodeling being a complex trait involving multiple genes (S5 Fig). Of note, the phenotypic spectra on ISO treatment were significantly broader than baseline, consistent with ISO perturbation enhancing the phenotypic variation among HMDP strains (Figure 3A and 3B). Moreover, across the HMDP strains, IVSd and FS increased significantly at week 1 and attenuated at later time points, reflecting the acute and chronic effects of ISO on septal wall hypertrophy and cardiac inotropy. LVIDd and LVM progressively increased with ISO treatment at each time point, reflecting progressive changes in LV chamber dimension and LV mass due to ISO (S4 Table). While many of the strains follow the general population trend described above, strains particularly susceptible to adverse cardiac remodeling, such as KK/HlJ and BTBRT $\times$ tf/J, demonstrated significantly decreased FS and increased in LVIDd (Figure 3C). The observed phenotypic variation, especially upon ISO treatment, underpins the basis for association mapping. The strain-level phenotype data are provided in S5 Table for reference.

### **The HMDP demonstrates the presence of a compensatory-decompensating phenotypic spectrum**

We performed correlation analysis across the HMDP to further understand the phenotype data observed. Not surprisingly, baseline body weight (BBW), a surrogate for body size, was significantly correlated with LVM and LVIDd but not with FS (S6A Fig), consistent with larger mice having larger and heavier hearts but body size not being a determinant of systolic heart function. In addition, under the baseline and ISO-treated conditions, HR was positively

correlated with FS and negatively correlated with LVIDd. This observation is consistent with known negative chronotropic and inotropic effects of isoflurane on the heart and reflects the varying degrees of isoflurane sensitivity among HMDP strain. In fact, HR was significantly correlated across control and different ISO-treated time points, indicative of a strain-specific HR response to ISO. (S6B and S6C Figs). Interestingly, HR had no significant correlation with LVM estimates. Based on these findings, we chose to correct for confounding effects of body size and isoflurane sensitivity using BBW and HR as surrogate markers in downstream analyses.

BBW- and HR-adjusted partial correlations among pairs of echocardiographic and LV weight traits showed that ISO treatment uncoupled many measures of cardiac structure and function from baseline (Figure 4A). ISO-treatment introduced an immediate perturbation on FS, such that FS from week 1 onwards was uncorrelated to FS at baseline, possibly due to different genetic factors controlling FS before and after ISO. In fact, FS at week 1 was negatively correlated with LVIDd at later time points, consistent with week 1 FS being a measure of contractile reserve and an early predictor of progressive ventricular dilation. In contrast, ISO-induced perturbation on IVSd was not observed until week 2. IVSd hypertrophy was positively correlated with increased FS and negatively correlated with LVIDd dilation, consistent with the presence of a compensatory-decompensating spectrum that is genetically controlled and required for association mapping (Figure 4B).

To describe each strain based on a composite of correlated traits, we used principal component (PC) analysis to transform ISO-induced changes into a set of linearly uncorrelated variables. The first three PCs account for 35%, 28% and 11% of the inter-strain variance observed, with the remaining PCs each accounting for < 10% of the variance observed. The variables factor map showed that the first three PCs captured phenotypes such as LVIDd dilation

and FS decrease; LVM hypertrophy; and early IVSd increase (Figure 5A). The individuals factor maps highlighted KK/HIJ with decompensating features, such as LVIDd dilation and FS decrease, and NZW/LacJ with compensatory features, such as late IVSd and LVM increase and preserved FS (Figure 5B and 5C).

### **Additive effects account for roughly half of the heritability estimates of cardiac structure and function across the HMDP**

Heritability is the proportion of observed differences due to genetic variation in the population. Specifically, broad-sense heritability ( $H^2$ ) reflects all the genetic contributions to a population's phenotypic variance including additive, dominant, and epistatic effects, while narrow-sense heritability ( $h^2$ ) represents the proportion of phenotypic variance that is due to additive genetic effects alone. In inbred model organisms, where the environment is controlled and multiple animals of identical genetic background are available, between-strain and within-strain variation can be used to estimate  $H^2$ , which reflects all genetic contribution to phenotypic variance. The majority of the traits we examined were highly heritable ( $H^2 > 0.4$ ) (S7 Fig).  $H^2$  estimates of LVM in the HMDP were between 61 and 81% (Table 1A), which is similar to estimates for other common traits, such as body fat and insulin resistance, in the HMDP panel (18) and consistent with genetic factors making a significant contribution to cardiac structure and function. Narrow-sense heritability  $h^2$ , the proportion of phenotypic variance due to additive genetic effects alone, incorporates genetic marker information and is the theoretical upper bound of what could be mapped in an association study. Narrow-sense heritability estimates  $h^2$  of LVM were from 39 to 57% (Table 1B). Our results suggest that additive and non-additive effects, such as dominant and epistatic gene x gene interactions, each account for roughly half of the broad-sense heritability estimates in cardiac structural and functional traits in mice.

## **ISO-induced gene expression changes across the HMDP are consistent with molecular, cellular and extracellular changes of cardiac remodeling**

Human heart failure is characterized by chronic sympathetic activation to compensate for progressive cardiac function loss. Chronically, high levels of catecholamines may lead to calcium overload, initiating a cascade of alterations at the cellular level to result in cardiomyocyte death and progressive deterioration in cardiac structure and function. To protect itself from harmful effects of chronic adrenergic stimulation, a protective mechanism, called desensitization, reduced membrane availability of the beta-adrenergic receptors, leading to diminished downstream signaling. ISO is a non-selective  $\beta$ -adrenergic agonist that, when administered chronically in mice, mimics the state of chronic sympathetic activation in human heart failure (41).

Consistent with molecular changes of cardiac remodeling in humans, ISO resulted in gene expression changes downstream of  $\beta$ -adrenergic receptor signaling across the HMDP mice.  $\beta$ 1-adrenoceptor expression decreased by 17% (*Adrb1*, adjusted p-value =  $2.56 \times 10^{-5}$ ) and the inhibitory G-protein  $G_i$  increased by 16% (*Gnai2*, adjusted p-value =  $4.47 \times 10^{-7}$ ), reminiscent of diminished  $\beta$ -adrenergic responsiveness secondary to chronic  $\beta$ -adrenergic overdrive in cardiomyopathic hearts (S2A Table). In addition, ISO downregulated *Myh7* expression by 34% and *Myh6* expression by 15% (adjusted p-value =  $6.93 \times 10^{-7}$  and  $6.77 \times 10^{-4}$ , respectively), possibly reflecting down-regulation of cardiac structural proteins or increased fibrotic tissue. Finally, one of the most significantly up-regulated genes *Lgals3* was increased by 3.5-fold (adjusted p-value =  $1.62 \times 10^{-22}$ ). *Lgals3*, or galectin-3, is an emerging clinical marker of cardiac fibrosis, adverse LV remodeling, and mortality in HF patients (42).

Gene ontology analysis of the up- and down-regulated probes was performed to examine the global changes in gene expression due to ISO. The up-regulated probes were most enriched for secreted signal glycoprotein, proteinaceous extracellular matrix (ECM), angiogenesis, polysaccharide binding, actin cytoskeleton, vacuole, response to wounding, chemokine signaling pathway, and epidermal growth factor (EGF)-like calcium-binding, prenylation and growth factor binding (S2B Table). The down-regulated probes were most enriched in mitochondrial matrix, mitochondrial inner membrane, and flavoprotein (S2C Table). Gene ontology analysis of the top 1000 correlated transcripts for each phenotype-expression pair was performed to identify processes specific for each phenotype. Under the baseline condition, transcripts that correlated with LVIDd were modestly enriched for actin-binding, cell junction and synapse. At week 3 of ISO, transcripts that correlated with LVID and LVM were highly enriched for proteinaceous ECM, secreted signal glycoprotein, actin cytoskeleton, collagen, ECM receptor interaction, EGF, polysaccharide binding, and mitochondrial membrane (S2D Table). These findings provided strong evidence that ISO induced cellular and ECM changes of cardiac remodeling found in HF.

### **Association mapping of cardiac remodeling and gene expression identified candidate genes for 17 cQTL associated with ISO-induced cardiac remodeling traits**

Association mapping at each ISO-treated time point was performed to define cQTLs for BBW- and HR-adjusted echocardiographic changes in measures from baseline. In addition, association mapping was performed for each gene transcripts under baseline and ISO-treated conditions, to define local expression quantitative trait loci (cis-eQTL) near each cQTL. Because genetic information flows from DNA to transcript to clinical phenotypes, a transcript with an eQTL that coincides with cQTL of a trait of interest may be hypothesized to be causal for the trait. Using a custom pipeline that overlays cQTL, cis-eQTL, expression-phenotype correlation,

and structural variation, candidate genes for each cQTL were prioritized. The significantly associated loci and candidate genes for echocardiographic measures are provided in Table 2. All significantly associated loci as well as supporting evidence described above are provided in S6 Table. All SNPs with association p-value < 0.001 are provided in S7 Table.

A total of 17 genome-wide significant cQTL were identified, including 4 for IVSd, 5 for LVIDd, 4 for LVM and 4 for FS, many of which were less than 1 Mb in size (Table 2). As expected, some of the correlated measures mapped to the same loci across ISO-treated time points or traits (S8 Fig). We compared our loci with available cardiovascular disease GWAS loci in human (S8 Table). One of the loci for week 1 IVSd hypertrophy near *Fam46a* on chromosome 9 has previously been shown in a multi-ethnic GWAS meta-analysis to be associated with systolic and diastolic blood pressure (p-value =  $6.2 \times 10^{-5}$  and  $5.5 \times 10^{-7}$ , respectively) (43). One of the loci for week 1 FS near *Lrp12* and *Zfp2* on chromosome 15 has previously been shown in a small human GWAS to be associated with sudden cardiac arrest due to ventricular tachycardia and ventricular fibrillation in patients with coronary artery disease (p-value =  $1 \times 10^{-6}$ ) (44).

### ***Klf4* is a causal gene for ISO-induced LVM hypertrophy**

A locus on chromosome 4 was significantly associated with LVM at week 3 (Figure 6A). LVM hypertrophy at weeks 1 and 2 and LV weight hypertrophy were also associated with this locus at lesser significance levels, while baseline LVM did not map to this locus (S8A-C Figs). Mouse strains with the GG and the TT genotypes at the lead SNP rs2779449 (p-value =  $9.3 \times 10^{-7}$ ) conferred a median week 3 LVM of 117 mg and 135 mg, respectively. SNP rs27794497, located in an intergenic region between *Tmem38b* and *Zfp462*, and surrounding SNPs in LD ( $r^2 > 0.8$ ) spanned a region containing 4 genes (Fig 6B). Among these genes, *Klf4* alone had a

significant cis-eQTL at SNP rs27794497 (p-value =  $5.9 \times 10^{-5}$ ) and was significantly correlated with LVM hypertrophy at week 3 (bikor = -0.3, p-value = 0.0088; S8D and S9 Figs). Our results suggest that *Klf4* is a negative modulator of ISO-induced LVM. During the course of this study, Yoshida *et al.* demonstrated that, when subjected to ISO, cardiomyocyte-specific *Klf4* knock-out mice demonstrated enhanced cardiac hypertrophy, augmented cellular enlargement, exaggerated expression of the transcriptional regulator myocardin and fetal cardiac genes (45).

***Myh14*** is a causal gene for ISO-induced LVM hypertrophy

In addition, a locus on chromosome 7 was significantly associated with LVM hypertrophy at weeks 3 (Figure 6A). This locus was also associated with LVM hypertrophy at week 1 and LV weight hypertrophy at lesser significance levels, while baseline LVM did not map to this locus (Figure 7). Mouse strains with the AA and the GG genotypes at the lead SNP rs40560913 (p-value =  $1.44 \times 10^{-9}$ ) conferred a median LVM increase of 43 mg and 28 mg, respectively. SNP rs40560913, located in the intron of *Izumo2*, and surrounding SNPs in LD ( $r^2 > 0.8$ ) spanned only one other gene *Myh14* (Figure 6C).

Neither *Izumo2*, a sperm-specific gene, nor *Myh14* was represented on the expression array used in our study. By qPCR of available matching control and ISO RNA samples in 26 mouse strains, we found *Myh14* expression to be increased by 46% with ISO treatment (p-value = 0.002). The available qPCR results was not adequate to map for eQTL. As *Myh14* was represented on microarray platforms used in prior HMDP studies, we queried our cis-eQTL database and found that *Myh14* has cis-eQTL signals across multiple tissues at the same locus and its expression in adipose and liver was negatively correlated with body mass normalized heart weight at week 8 in the HMDP obesity study (S9 and S10 Tables) (18). Based on these findings, we hypothesized that *Myh14* was the causal gene for this locus.

*Myh14* is a non-muscle myosin involved in mechanotransduction—knockout of the gene/protein has no basal phenotype [2] and the role of MYH14 has never been comprehensively determined in the heart. To validate *Myh14* as a causal gene for LVM hypertrophy, we performed siRNA knock-down of *Myh14* in neonatal rat ventricular myocytes (NRVM) followed by phenylephrine and ISO stimulation *in vitro*. *Myh14* knock-down demonstrated > 70% efficiency (Figure 8A). *Myh14* siRNA treated cells displayed poor attachment to the tissue culture plates (Figure 8B). In addition, quantitative measurements of cell size showed that *Myh14* siRNA treated NRVMs failed to undergo characteristic cellular hypertrophy under phenylephrine (PE) and ISO stimulation (Figure 8C) and resulted in decreased cell viability (Figure 8D).

To further examine the role of *Myh14* in cardiac remodeling *in vivo*, we obtained *Myh14*<sup>-/-</sup> mice from the Adelstein lab (40). Wild-type, heterozygous (*Myh14*<sup>+/-</sup>) and homozygous (*Myh14*<sup>-/-</sup>) littermates were not phenotypically different and no disproportionate spontaneous sudden death was observed in *Myh14* deficient mice (oldest *Myh14*<sup>-/-</sup> mouse 12 months to date). qPCR analysis showed a complete loss of *Myh14* expression in *Myh14*<sup>-/-</sup> hearts (Figure 9H). ISO-induced LVM and LVIDd hypertrophy were significantly increased in *Myh14*<sup>-/-</sup> and *Myh14*<sup>+/-</sup> versus wild-type controls (Figure 9A and 9B). There were no significant differences in baseline EF among genotypes but ISO-treated *Myh14*<sup>-/-</sup> mice demonstrated a trend towards EF reduction compared to *Myh14*<sup>+/-</sup> and WT mice (Figure 9C). Cardiomyocyte cross-sectional area, as measured by wheat germ agglutinin (WGA) staining, showed that cardiomyocytes from ISO-treated *Myh14*<sup>-/-</sup> mice were more significantly hypertrophied compared to *Myh14*<sup>+/-</sup> and WT mice (Figure 9D and 9E). Histological sections of *Myh14*<sup>-/-</sup> hearts revealed no abnormal phenotype at baseline but increased myocardial fibrosis and intercalated disc disarray after ISO treatment

(Figure 9F and 9G). RT-PCR demonstrated the up-regulation of immediate early gene *Myc*, fetal gene *Nppb*, and fibrosis gene *Lgals3* in ISO-treated *Myh14* deficient hearts compared to controls (Figure 9H).

### **The HMDP Resource**

The study data, including gene-gene and gene-trait correlations as well as clinical trait and transcript mapping, can be accessed via a user-friendly web-based interface at

<http://systems.genetics.ucla.edu/>.

## Discussion

Human HF GWAS have identified a handful of genetic determinants that explain a small proportion of the genetic variance, in spite of involving tens of thousands of individuals (46). Given the breadth of etiological and clinico-pathological heterogeneity, even larger cohorts may not be big enough to discriminate signals from noise. We carried out an alternative approach in the mouse, which has several advantages over human. First, we used multiple mice per strain and were able to fully control the age, environment, severity and timing of cardiac injury, to accurately assess cardiac remodeling with minimal confounding. As a result, our heritability estimates of LVM were significantly higher than that in human of 24-32% in the Framingham Heart Study (47) and 36-47% in an Australian twin study (48). The relatively high heritability provided sufficient power to uncover novel loci not previously identified in human. Second, the HMDP strains have either been sequenced or densely genotyped, eliminating genotyping costs and errors. Third, access to control and ISO LV transcriptome as well as sequence variation in 17 of the classical inbred strains available from the Wellcome Trust Mouse Genomes Project (MGP) allowed us to examine cis-eQTL, transcript-to-trait correlations, and functional variants to prioritize candidate genes and generate hypotheses for functional validation without prior knowledge in an unbiased manner, thus overcoming some of the challenges of working with relatively long LD in mice. Given the paucity of association signals in humans, we consider the association signals we observed in mice to be complementary to existing human studies. While the individual genetic variants observed in mice may not be identical to the variants observed in human populations, candidate genes from this study help identify genes and/or pathways participating in modifying mechanisms of heart failure pathology.

We recognize that using only one gender is a limitation of our study that prevents us to readily extrapolate our findings to male mice. During the pilot phase of the study, we observed a slightly wider phenotypic separation among the 4 core classical inbred strains (A/J, C57BL/6J, C3H/HeJ, and DBA/2J) in females than in males, which is important for our association mapping study design. Under the constraints of cost for animal procurement and maintenance, we chose to study females only. We believe that at least some of the findings observed in females will likely recapitulate in males, since gender differences in mortality in humans appear to be predominantly due to factors, such as age, co-morbidities and treatment approaches (49), rather than gender itself. Moreover, higher rates of mortality and cardiogenic shock have been observed in women compared with men after acute myocardial infarction (49), yet females are under represented in both human and mouse studies.

Our study established an echocardiographic reference range for 100+ inbred mouse strains. To minimize the variability to handling among different mouse strains, inhaled isoflurane, one of the most commonly used agent for prolonged anesthesia in mice, was administered during echocardiography. Based on recent studies on sedated mice, FS of 33% and 32% in 129X1/SvJ male and C57BL/6J female mice were similar to our strain averages of 33% and 34%, respectively (50, 51).

Inhaled anesthetics, including isoflurane, are known cardiac depressants, especially upon deep sedation. In a study involving C57BL/6 male mice, mean arterial pressure (MAP), HR, and body temperature at an isoflurane concentrations of 1.5% (volume-to-volume, v/v) remained comparable to the conscious state for up to 90 minutes post-induction (52). In another study using C57BL/6 male mice, where HR was fixed at different dosages/depths of isoflurane (53), echocardiographic measurements were found to be highly reproducible at a high HR (475–525

bpm) but were correlated with HR under a low HR (350–400 bpm). Although no comparable prior data exist for the vast majority of the mouse strains in our study, we empirically observed that different mouse strains exhibited variable sensitivity to anesthesia and that HR correlated to a number of echocardiographic measures representing cardiac inotropy and chamber size (Fig S6B). Therefore, we adjusted for HR as a covariate and a surrogate for isoflurane sensitivity in our analysis. In spite of concerns regarding HR being a confounding factor and surrogate marker for variable isoflurane sensitivity, correlations among echocardiographic measures remained roughly unchanged, before and after adjusting for HR, consistent with relatively minor cardiac depressive effects of isoflurane.

In spite of ongoing inotropic stimulation by ISO throughout the 3-week treatment time course, some strains displayed the same or decreased FS at later ISO time points compared to baseline, consistent with a decrease in contractile reserve -- an observation in early HF stages of decreased cardiac function augmentation to direct inotropic stimulation -- marked by reduced beta-adrenergic receptors along with altered sarcoplasmic reticulum Ca<sup>2+</sup>-adenosine triphosphatase 2a (SERCA2a) and phospholamban (54). Our narrow-sense heritability estimates were roughly half of broad-sense heritability estimates, suggesting that both additive and dominant/epistatic genetic effects play important roles in the genetic variance of HF.

Our association analyses provided direct evidence that common genetic variation control ISO-induced cardiac remodeling. The following considerations are relevant to the interpretation of the results. Equivalent dosages of ISO were given to each strain to control the degree of cardiac insult. Although we could not control for ISO metabolism, we believe its contribution to the results was minor. Linkage analysis using only the BXD RI strains resulted in a significant reduction in resolution (S10 Fig). On the other hand, association analysis using only the classical

inbred strains suffered a significant loss of power (S11 Fig). Our study, which combines the classical inbred and RI strains to achieve high genetic resolution and adequate statistical power, represents a substantial improvement over traditional linkage analysis. In addition, panel-specific SNP selection based on minor allele frequency and missing genotype cutoffs could indirectly affect association results. For example, SNPs that did not meet the minor allele frequency and missing genotype cutoffs in the entire HMDP panel resulted in the loss of association signals around the 40 Mb region on chromosome 7 (S11C Fig). Moreover, genetic variation that is well represented among the BXD panel could contribute more substantially to the association results, due to the fact that the BXD panel (44 strains) was highly represented among the entire HMDP panel (104 strains) (S10 Fig). In the setting of a relatively narrow phenotypic spectrum, adjustments for uncorrelated covariates can deflate the association results. For example, baseline FS, which was not correlated to BBW and has a similar narrow-sense heritability estimate as ISO-treated FS, had two associated loci prior to but none after BBW adjustment. Finally, the significantly associated loci explained about 23-38% of the total variance observed, which is significantly higher than in human HF GWAS.

In our previously published manuscript, we identified a number of heart weight and fibrosis loci but none for ISO-induced LV weight hypertrophy (21). A number of factors may have hampered the detection of association signals for LV weight. For example, LV weight hypertrophy for each strain was estimated based on the difference of the strain averages from control and ISO-treated cohorts. Given the small numbers of individuals in each cohort, variation in LV weight due to variation in body size and cardiac chamber dissection may have introduced errors in ISO-induced LV weight hypertrophy estimates per strain. In contrast, LVM was measured non-invasively, which allowed for repeated measurements in the same animal. The

baseline LVM estimate served as an internal control measure for body size in each mouse treated with ISO. This approach represented an important advantage over our previous analysis, which allowed us to estimate ISO-induced LVM hypertrophy more accurately and improve the power of association analysis. Of note, the LVM measurements demonstrated a positive bias in the upper ranges as a result of its derivation, which contains a constant multiplier:  $LV\ mass\ (Penn) = 1.04 ([LVIDd + PWd + IVSd]^3 - [LVIDd]^3) - 13,6\ g$ , where LVIDd represents left ventricular internal dimension at end diastole, PWd represents posterior wall thickness at end diastole, and IVSd represents interventricular septal wall thickness at end diastole.

In a mapping study, the spread rather than the absolute values of the data drives the association analysis. Therefore, in spite of the positive bias in the upper ranges, LVM and LV weights yielded similar association results (Fig 6C and Fig 7D), with LV weight having a reduced statistical significance as compared to LVM.

Our association analysis not only highlighted a number of genomic regions and genes for HF susceptibility in mice, it also provides independent supportive evidence for syntenic regions in humans and is complementary to existing and future human HF genome-wide association and family-based linkage studies. For example, a candidate causal genomic region or gene for HF or inherited cardiomyopathy in humans not reaching statistical significance may be recovered based on additional supportive evidence from mice. Next, based on the central dogma that the flow of genomic information starts from the DNA, to the RNA then to the phenotype of interest, the cardiac transcriptome data allow us to generate hypotheses to test causal inference. For instance, when the transcript level of a gene in an association locus is associated with genetic variation and correlates with phenotype, we may hypothesize that the gene could be causally related to the phenotype, thereby facilitating prioritization of candidate genes in an association locus (Fig 10).

Our data could be used to explore the roles of individual candidate genes in existing human GWAS loci. In addition, we have provided a rich data resource for the HF research community to interrogate a given gene's relationship with other genes in the cardiac transcriptome. For example, a HF community researcher studying a novel HF gene X may be interested in its upstream regulatory and downstream regulated genes. In addition to identifying other genes in the cardiac transcriptome that correlated with gene X, a search for gene X's trans-eQTL, a distant locus controlling the expression of gene X, may reveal a number of putative upstream regulators of gene X for hypothesis driven experiments and testing. On the other hand, if gene X was regulated at the local level, identification of additional cardiac transcripts regulated by the same locus could reveal gene X's putative downstream targets. Finally, the differentially expressed cardiac transcripts that correlated with cardiac remodeling phenotypes represent a set of genes that may play important roles in HF pathogenesis and compensatory changes and may be enriched for HF biomarkers. Our transcriptome data could complement existing human transcriptome data, a valuable set of transcripts that represent the true spectrum of human HF but whose RNA sample quality may be more variable, to identify additional players of HF pathogenesis, compensatory changes and biomarkers.

Our study implicated *Klf4* and *Myh14* to be negative regulators of cardiac hypertrophy under ISO stress. Our results showed that the degree of LVM hypertrophy was inversely correlated with the expression of *Klf4* (S9 Fig). Since the expression of *Klf4* expression was genetically determined by the cis-eQTL, our findings suggest that *Klf4* expression may be causally related to the degree of LVM hypertrophy. In other words, the higher the *Klf4* expression, the lesser the degree of LVM hypertrophy. We note, however, that our data does not provide proof of a causal relationship, since a significant fraction of *Klf4* expression occurs in

trans, but is consistent with findings from previous literature showing that KLF4 is a negative modulator of ISO-induced LVM. *Klf4*, or Krüppel-like factor 4, a zinc-finger transcriptional regulator best known as one of the four Yamanaka factors that are sufficient to reprogram differentiated cells into embryonic-like induced-pluripotent stem cells (iPSCs) (55), has recently emerged as an important modulator of cardiomyocyte hypertrophy. Cardiomyocyte-specific deletion of *Klf4* in mice resulted in a slightly increased heart weight and increased ANF levels compared to controls. When subjected to pressure overload, these mice developed increased pathologic cardiac hypertrophy, fibrosis, apoptosis and mortality, compared to control banded mice (56). *In vitro* studies confirmed that KLF4 binds to the ANF and GATA4 promoters and acts as a repressor of cardiomyocyte hypertrophy (57).

*Myh14* was the only gene expressed in the heart that was located at the chromosome 7 locus affecting cardiac hypertrophy in our study. *Myh14*, myosin heavy polypeptide 14, encodes the heavy chain of the molecular motor nonmuscle myosin II-C (NMIIC), which is a member of the nonmuscle myosin II motor protein family that plays an integral role in mechanotransduction, converging external and cell-generated forces by interacting with cytoskeletal actin. When bred in a background of non-muscle myosin II-B (NMIIB) hypomorphic mice expressing only 12% of wild-type amounts of NMIIB protein, *Myh14*<sup>-/-</sup> mice developed marked cardiac myocyte hypertrophy, interstitial fibrosis and diffuse N-cadherin and  $\beta$ -catenin patterns at the intercalated discs, where NMIIB and NMIIC colocalized (40). Collectively, these prior studies implicated *Klf4* and *Myh14* as negative regulators of stress-induced cardiac hypertrophy *in vivo*, which is also supported by our results.

In our *in vivo* study, we showed that *Myh14* deficient mice exhibited increased ISO-induced LVID dilation, LVM hypertrophy, cardiac fibrosis, and hypertrophic markers *Myc* and

*Nppb*. On the other hand, *In vitro* siRNA knockdown of *Myh14* in NRVMs failed to elicit a hypertrophic response when stimulated with ISO but rather displayed poor attachment to the tissue culture plates and poor survival. As a non-muscle myosin, *Myh14* is ubiquitously expressed and could act through non-cardiomyocytes to modulate cardiomyocyte growth and hypertrophy. We focused on *Myh14* inside cardiomyocytes due to a previous study demonstrating the location of *Myh14* in the intercalated disc and its implicated role at the cell-cell junction (40). Since NRVMs in tissue culture do not fully recapitulate the three-dimensional structure of the heart, the consequence of *Myh14* deficiency may be quite different *in vitro* versus *in vivo*. These seemingly discrepant results raise important questions regarding the mechanisms through which *Myh14* deficiency leads to cardiomyocyte death in the absence of mechanotransduction stress *in vitro* and cardiac hypertrophy on an organ level *in vivo*.

To further explore mechanisms mediating *Myh14* deficiency in LVM hypertrophy, we queried the existing HMDP cardiac transcriptome database available via <http://systems.genetics.ucla.edu/> (58). We found that the cis-eQTL locus that controls *Myh14* expression also controls *Foxo1* in trans (p-value =  $1.20 \times 10^{-5}$ ) and that the expression of *Foxo1* was significantly positively correlated with *Myh14* (bicor = 0.74, p-value =  $6.7 \times 10^{-18}$ ), which implicated *Foxo1*, a forkhead family transcription factor, as a potential mediator of *Myh14*. In addition, *Myh14* expression was correlated with *Cdk11b* (bicor = 0.59, p-value  $1.53 \times 10^{-10}$ ) and *Cdk11b* was controlled by the *Myh14* locus (rs32804715, p-value =  $9.08 \times 10^{-6}$ ), which implicated *Cdk11b*, a member of the cyclin-dependent kinase family known to be a negative regulator of normal cell cycle progression, as a potential mediator of *Myh14*.

Moreover, we showed that ISO-treated *Myh14* deficient heart demonstrated disrupted assembly of intercalated discs involving  $\beta$ -catenin, which is a dual function protein that provides

cell-cell connections at the adherens junctions in the intercalated disc, to facilitate mechanical and electrical coupling between adjacent cardiomyocytes, and regulates gene transcription as a transcriptional coactivator, raising the possibility that disordered  $\beta$ -catenin localization in the intercalated disc due to *Myh14* deficiency could have direct consequences on  $\beta$ -catenin signaling. In fact, altered intercalated disc architecture in cardiac muscle hypertrophy and HF has previously been associated with altered  $\beta$ -catenin abundance (59), subcellular localization (60), and canonical Wnt/ $\beta$ -catenin/TCF signaling. Interestingly, transcription factors FOXO and TCF are known to compete to interact with  $\beta$ -catenin (61). Our experimental results showed that *Myh14* is important in the maintenance of normal cardiac structure under ISO stress and that the loss of *Myh14* exacerbates the hypertrophic response to ISO. Further understanding of the crosstalk between *Myh14*,  $\beta$ -catenin, and *Foxo1*, will likely reveal important mechanisms regarding how *Myh14* deficiency leads to cardiac hypertrophy under ISO stress.

Cardiac remodeling is one of the most important prognostic determinants of clinical HF. Our study results provide an important resource to the HF research community and highlight the strength of a systems approach to studying HF. The integration of high throughput molecular phenotypes, such as genomic and transcriptomic data, provides a means to identify novel candidate causal genes and a powerful alternative to human studies to understand the complex interactions underpinning phenotypic variation of cardiac remodeling. Similar to human GWAS, many of our lead SNPs lie in noncoding regions. Future directions will include sequence-based computational approaches to systematically prioritize functional regulatory variants. Future insights in how common genetic variations in a population modify HF progression will further shed light on genetic risk profiling, gene-gene and gene-environment interactions as well as the design of personalized therapies for HF patients.

## **Acknowledgments**

Drs. Robert Adelstein and Xuefei Ma kindly provided the Myh14 knockout mice. Profs Tom Vondriska and Eleazar Eskin contributed to critical review of the manuscript.

**Table 1. Heritability estimates of cardiac structure and function under isoproterenol treatment.**

A. Broad-sense heritability estimates based on line repeatability.

$r^2$	Control	Week 1	Week 2	Week 3
IVSd	0.61	0.59	0.49	0.43
LVIDd	0.76	0.68	0.67	0.68
LVM	0.81	0.70	0.61	0.74
FS	0.73	0.64	0.60	0.68

B. Narrow-sense heritability estimates based on marker-based heritability estimates.

$r^2$	Control	Week 1	Week 2	Week 3
IVSd	0.20 (0.12-0.31)	0.27 (0.15-0.44)	0.21 (0.09-0.41)	0.15 (0.06-0.35)
LVIDd	0.40 (0.29-0.51)	0.41 (0.27-0.57)	0.45 (0.31-0.60)	0.49 (0.34-0.64)
LVM	0.48 (0.38-0.59)	0.44 (0.31-0.58)	0.39 (0.26-0.55)	0.57 (0.43-0.70)
FS	0.30 (0.21-0.41)	0.31 (0.18-0.47)	0.34 (0.20-0.50)	0.38 (0.24-0.55)

Ninety-five percent confidence intervals are represented within the parentheses.

**Table 2. Significant association loci.** Start and end positions as well as range are in the unit of Mb. Peak p-value and SNP rsID and are listed. If multiple SNPs have the same p-value, only one representative rsID is listed and indicated by an underline. MAF denotes the minor allele frequency. ES denotes the effect size per allele. The units for the effect sizes are mm for IVSd and LVIDd, mg for LVM, and % for FS. Variance explained is denoted by  $\omega^2$ . Analyses are the analyses in which the SNP exceeded the genome-wide significant threshold. If multiple analyses yielded significant p-value at a given SNP, the underlined analyses has the most statistically significant p-value. Candidate genes are denoted by structural variation (underlined), cis-eQTL (italics) and correlation with trait were present (bold). *Golph3l* and *Bmp5* have a stop\_gained variant. *M5C1000I18Rik* and *Dpys* have a splice acceptor variant. *Dpys* has 2 missense variants.

\* denotes that one of the lead SNPs resided in the gene

Chr	Start	End	Range	Trait	rsID	P-value	MAF	ES	$\omega^2$	Analyses	Candidate genes
1	187.4	187.6	0.2	LVIDd	rs32292745	2.97E-06	0.08	0.21	1%	1dLVIDd	<i>Lyplal1</i>
3	82.1	82.6	0.5	FS	rs32712632	9.06E-07	0.3	6.1	14%	3FS	Tdo2, Accn5, Gucy1a3
3	94.3	95.7	1.4	LVIDd	rs33064660	1.30E-06	0.22	0.29	7%	3LVIDd	<i>Rorc</i> , <i>Oaz3</i> , <i>Mrp19</i> , <i>Snx27*</i> , <i>Selenbp1</i> , <i>Psm4</i> , <i>Tmod4</i> , <i>Gm128</i> , <i>Bnpl</i> , <i>Anxa9</i> , <i>Lass2</i> , <i>Ctsk</i> , <i>Golph3l</i> , <i>Ecm1</i> , <i>Prpf3</i> , <i>Mrps21</i> , <i>Gm129</i> , <i>BC028528</i> , <i>Car14</i>
4	54.2	55	0.8	LVM	rs27794497	9.30E-07	0.36	18.2	11%	<u>3LVM</u> ,3dLVM,1LVM,1dLVM	<i>Klf4</i>
4	58.2	58.3	0.1	LVIDd	rs27851114	2.78E-06	0.06	0.2	6%	1dLVIDd	<i>Txndc8</i> , <i>Svep1</i> , <i>Musk</i>
4	62.4	62.7	0.3	LVM	rs28295600	3.77E-06	0.39	18.2	9%	3LVM	<i>Rgs3</i> , <i>Zfp618</i> , <i>Orm3</i> , <i>Whn</i>
4	63.7	64	0.3	LVM	<u>rs3656076</u>	1.74E-07	0.44	20.8	8%	<u>3LVM</u> ,3dLVM	<i>Tnc</i>
4	93.1	95.2	2.1	IVSd	<u>rs28128253</u>	4.65E-09	0.36	0.10	15%	1IVSd	<b>Jun</b>
7	51.5	52.4	0.9	LVM	rs40560913	1.44E-09	0.4	18.8	10%	3LVM,3dLVM,2LVIDd,iLV,dLV	<i>Myh14</i> , <i>Izumo2*</i>
9	66.9	69.6	2.7	IVSd	rs49424819	7.35E-07	0.35	0.09	1%	1dIVSd	<u>M5C1000I18Rik</u> , <i>Vps13c</i>
9	74.9	75.8	0.9	FS	rs33896682	3.41E-06	0.40	3.22	4%	2dFS	<i>Arpp19*</i> , <u><i>Bmp5</i></u>
9	75.8	80.2	4.4	IVSd	<u>rs13480288</u>	1.95E-09	0.38	0.10	12%	1dIVSd	<i>Tinag</i> , <i>Lrrc1*</i> , <i>Omt2b</i> , <i>Impg1</i>
9	84	84.6	0.6	IVSd	rs36266287	6.89E-07	0.38	0.09	6%	1dIVSd	<i>Bckdhib</i> , <i>Fam46a</i>
11	36.8	36.9	0.1	LVIDd	<u>rs6333970</u>	2.83E-06	0.31	0.25	8%	1LVIDd	<i>Tenn2</i>
12	57.8	57.9	0.1	LVIDd	rs47048438	3.41E-06	0.07	0.20	1%	1dLVIDd	<i>Nkx2-9</i> , <i>Slc25a21*</i>
15	40	40.1	0.1	FS	rs48791248	2.71E-07	0.10	6.79	14%	1FS	<u><i>Dpys</i></u> , <i>Lrp12</i> , <i>Zfpm2</i>
18	47	49.2	2.2	FS	<u>rs51860788</u>	1.04E-06	0.41	6.62	6%	2FS	<b>Trim36</b>

## **Figure Legends**

**Figure 1. Experimental Design.** We characterized 104 classical and recombinant inbred mouse strains of the Hybrid Mouse Diversity Panel (HMDP) by serial echocardiograms at baseline, week 1, week 2, and week 3 under the control condition or chronic isoproterenol (ISO) infusion. The  symbol indicates the time points echocardiograms were performed.

**Figure 2. Variation in echocardiographic measures of cardiac structure and function among HMDP mouse strains.** Black bars represent measurements under the baseline condition in ranked order. White bars represent measurements after 3 weeks of continuous ISO infusion. IVSd = interventricular septal wall thickness; LVIDd = left ventricular diastolic diameter; LVM = left ventricular mass; FS = fractional shortening. Error bars represent the standard errors of the means.

**Figure 3. Population distribution of echocardiographic measures at each time point.** The violin plot is a combination of a boxplot and a kernel density plot (a smooth histogram) rotated on its side. The white dot represents the median. The black box represents the interquartile range (IQR). The black vertical line represents the whiskers spanning the lowest and the highest data within 1.5 IQR from the lower and upper quartile. A) The population distribution at each ISO treatment time point. B) The distribution of the changes in echocardiographic measures from baseline. C) The changes in echocardiographic measures compared to baseline at each ISO time point for individual classical inbred strains.

**Figure 4. Correlations among echocardiographic traits across time points.** A) Correlations among baseline body weight- and heart rate-adjusted traits across time points. B) Correlations among isoproterenol-induced changes in traits from baseline across time points. Colored dots

represent the pairwise correlation  $r$  values with  $p$ -values exceeding Bonferroni corrected  $\alpha$  significance level of 0.05. LV represents LV weight. Suffixes represent weekly echocardiographic time points under ISO treatment (0 = baseline, 1 = week 1, 2 = week 2, 3 = week 3) or control versus ISO LV weight at week 3 (c = control, i = ISO).

**Figure 5. Principal component analysis of isoproterenol-induced changes from baseline. A.**

The variables factor map demonstrates the first three principal components that account for 74.8% of inter-strain variations. The first principal component corresponds roughly to LVIDd dilation and FS decrease. The second principal component corresponds to IVSd and LVM hypertrophy. The third principal component corresponds to early IVSd hypertrophy. The individual factor map projects the inbred mouse strains onto the first three principal components. Black dots highlight the classical (B) and recombinant (C) inbred strains and gray dots represent the remaining HMDP strains.

**Figure 6. Genome-wide association for week 3 LVM. A.** Manhattan plot for week 3 LVM. **B.** Regional plot for week 3 LVM around peak SNP rs27794497 (purple). **C.** Regional plot for week 3 LVM hypertrophy around SNP rs40560913 (purple). Pairwise  $r^2$  between the peak SNP and the surrounding SNPs are denoted by color scale.

**Figure 7. Regional plots of LV hypertrophy at chromosome 7 across time points. A.**

Regional plot for LVM at baseline. Regional plots for ISO-induced LVM hypertrophy at week 1 (B) and week 2 (C). **D.** Regional plot for ISO-induced LV weight hypertrophy at week 3.

**Figure 8. *Myh14* knock-down in stressed neonatal rat ventricular myocytes (NRVMs)**

**causes abrogation of hypertrophic response and decreased cell viability.** Knock-down of *Myh14* in NRVMs using siRNA (A), followed by ISO and phenylephrine (PHE) treatments, showed the *Myh14* knock-down treated NRVMs displayed poor attachment to tissue culture

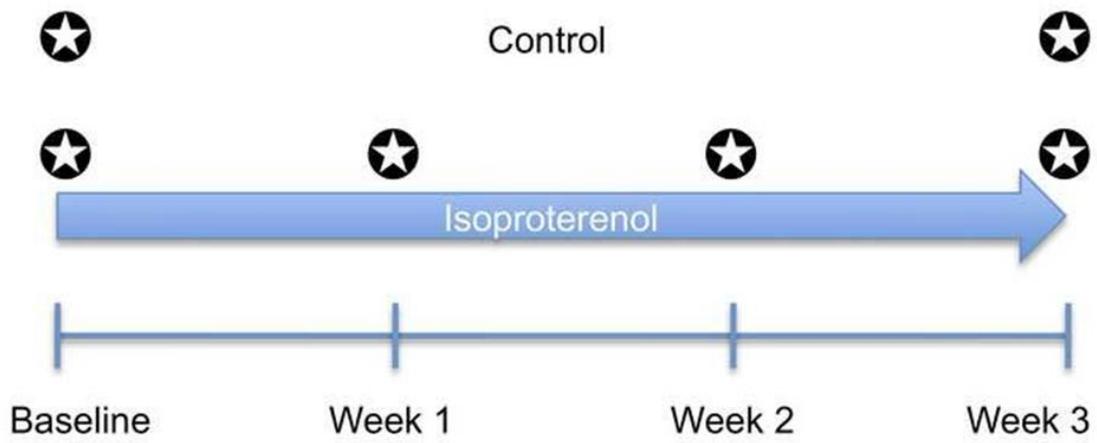
plates (B), did not undergo characteristic hypertrophy (C) and resulted in decreased cell viability based on MTT assay (D). Error bars are SEM, n=3. \* denotes  $p < 0.05$ .

**Figure 9. *Myh14* deficiency under ISO results in LVM hypertrophy, LVIDd dilation, increased cardiomyocyte size, fibrosis, intercalated disc disarray, and hypertrophic signals.**

Female mice of wild-type (WT), heterozygous (HET) and knockout (KO) *Myh14* genotypes were treated with ISO at 10-12 weeks of age (n = 6, 9, and 7, respectively). *Myh14* deficiency conferred an increase in ISO-induced LVM hypertrophy (A) and LVID dilation (B). There was a trend towards decreased ISO-induced EF (C) due to *Myh14* deficiency. Cardiomyocyte cross-sectional area, as measured by wheat germ agglutinin (WGA) staining (D), showed that cardiomyocytes from ISO-treated KO mice were more significantly hypertrophied compared to HET and WT mice (E). Dark gray bars (baseline in A and B; average of baseline and week 1 in C; control in E). Light gray bars (average of weeks 1-3 measures in A and B; average of weeks 2-3 measures in C; ISO-treated in E). *Myh14* deficiency conferred an increase in ISO-induced fibrosis (blue arrow) by Masson's trichrome staining (F) and an increase in intercalated disc disarray (blue arrow) by  $\beta$ -catenin staining (G). Cardiac tissue gene expression of hypertrophic and fibrosis markers were examined by RT-PCR at the end of a 3-week ISO infusion (H). *Myc*, *Nppb*, and *Lgals3* were increased with *Myh14* deficiency. Error bars are SEM, n=3. \* represents t-test p-value  $< 0.05$ . \*\* represents t-test p-value  $< 0.005$ .

**Figure 10. Relationships between variation and correlated traits.** There are three possible causal relationships when there is correlation among a SNP (single nucleotide polymorphism) S, a transcript (RNA) R, and a physiological or pathologic trait (phenotype) P. In the causal model, the SNP variation affects its transcript levels leading to the resulting phenotype. In the reactive

model, the SNP acts on the phenotypes, which in turn affects transcript. In the independent model, the SNP variation acts on both the phenotype and transcript independent.



**Figure 1.**



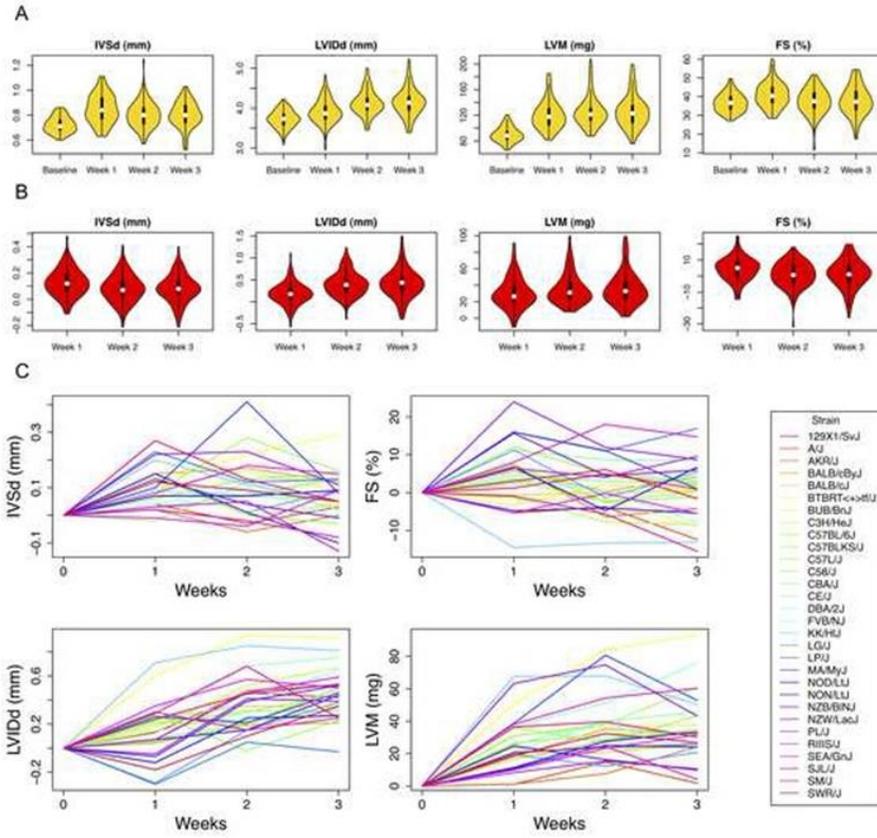
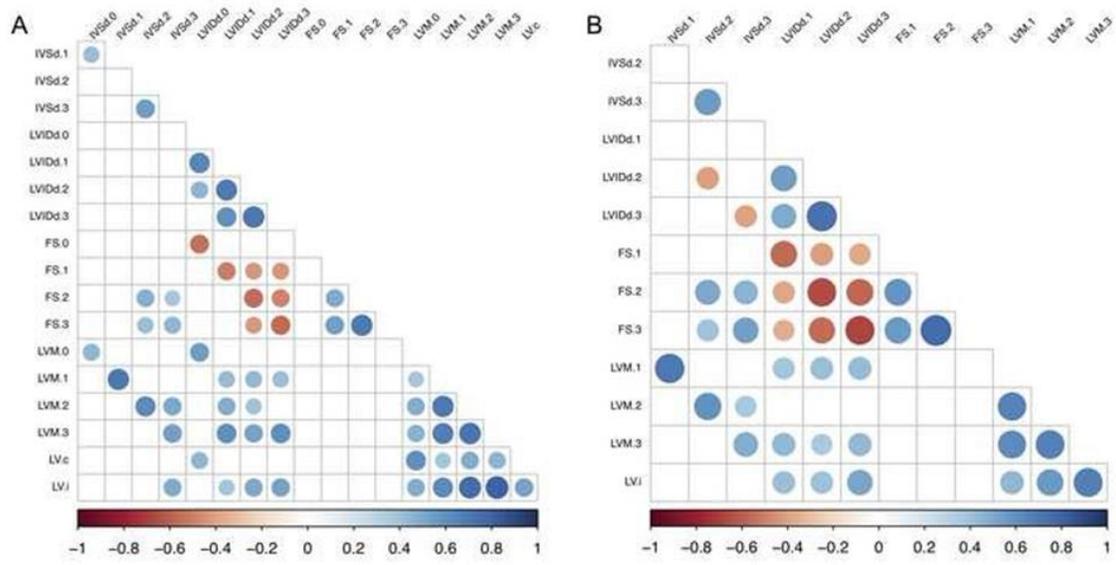


Figure 3.



**Figure 4.**

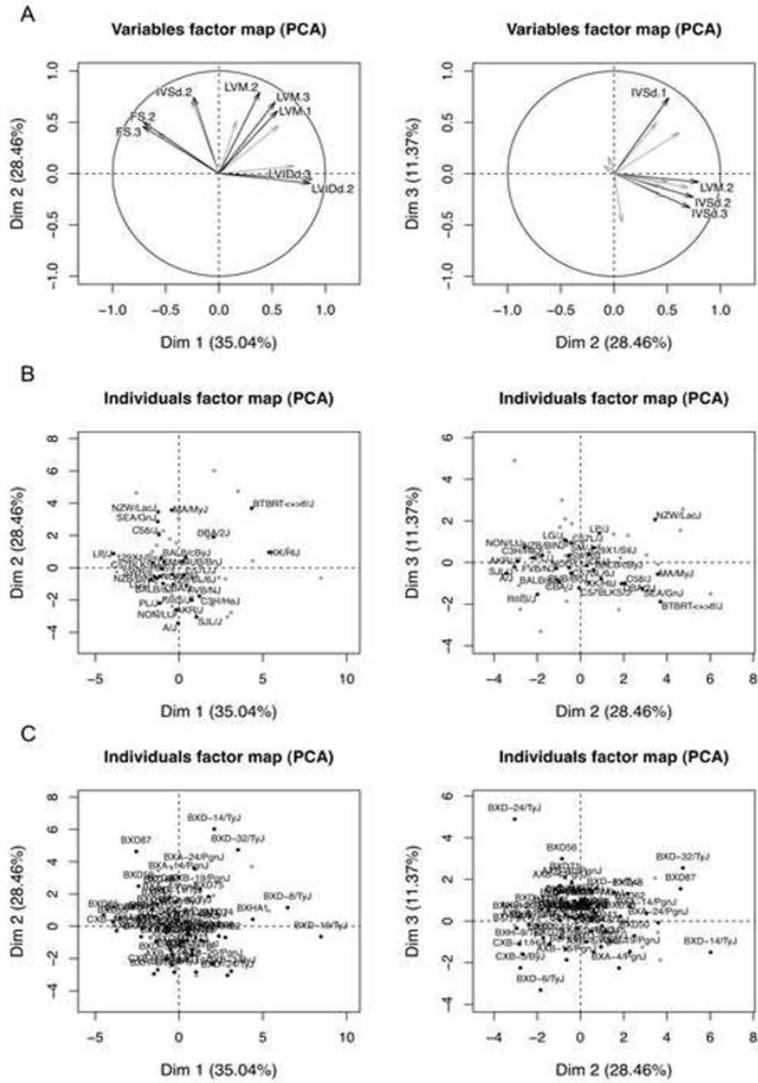
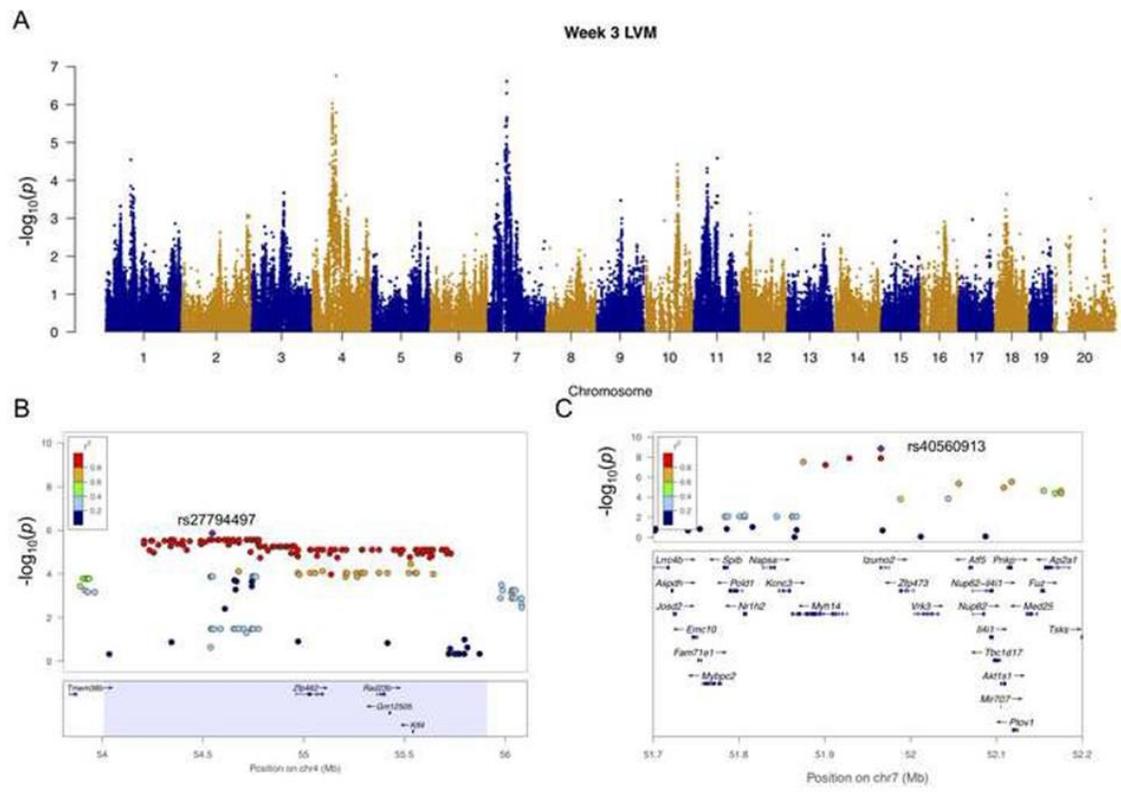


Figure 5.



**Figure 6.**

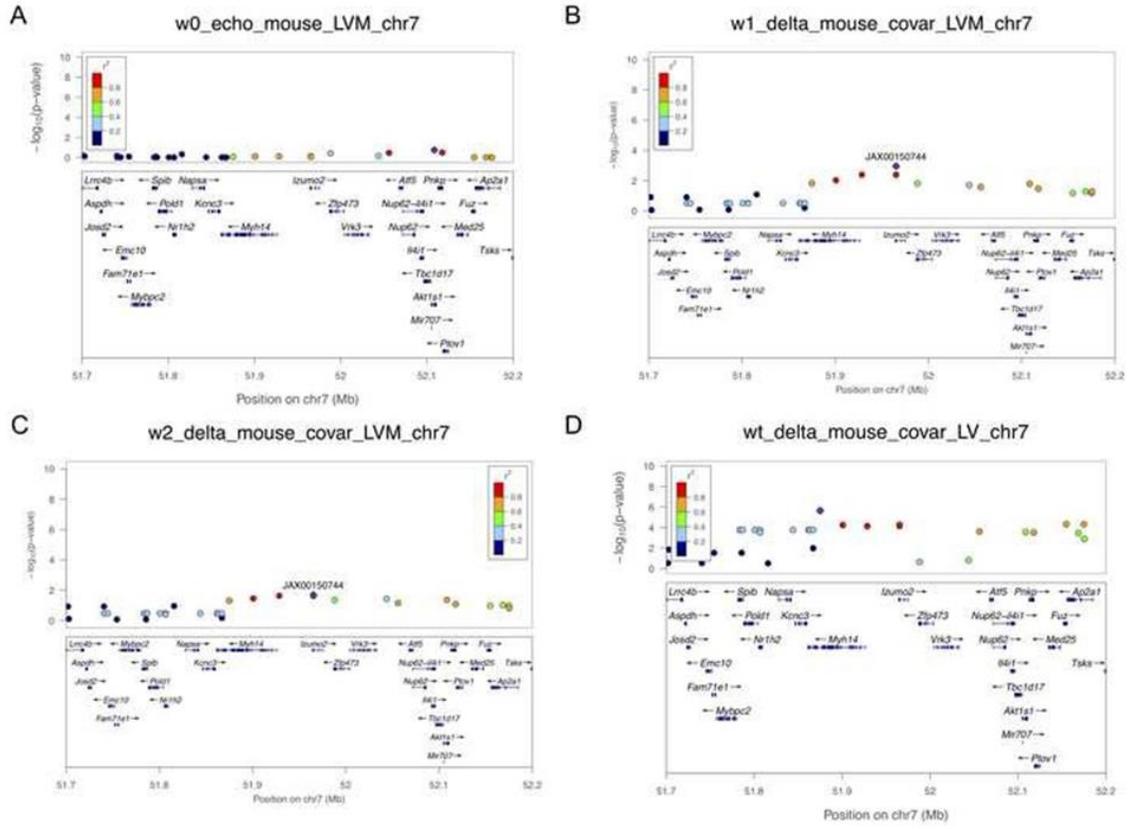
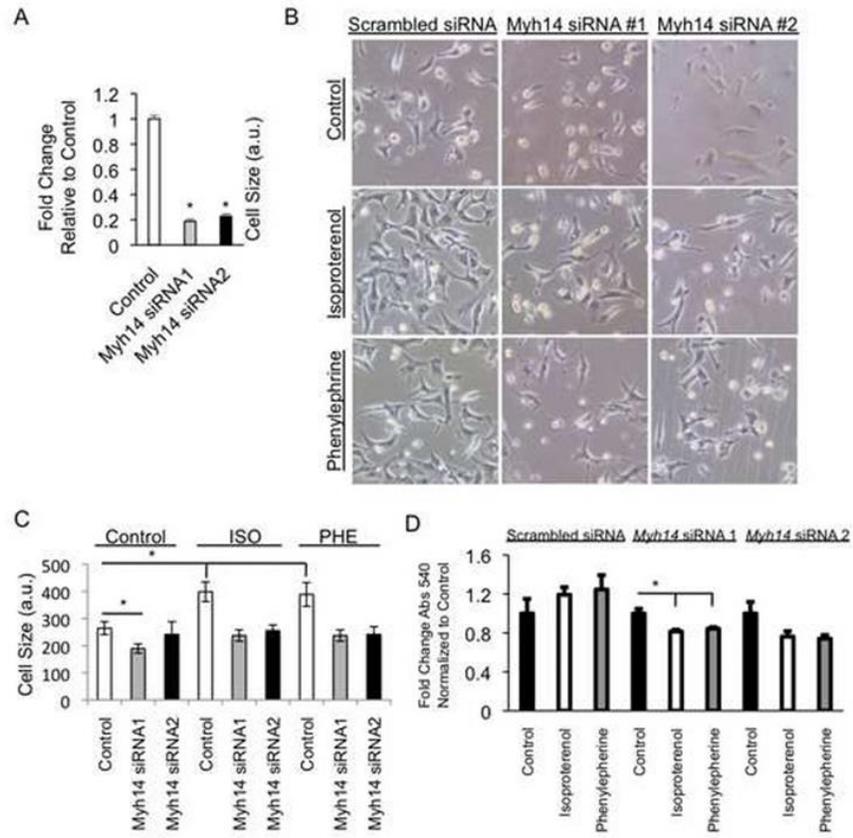
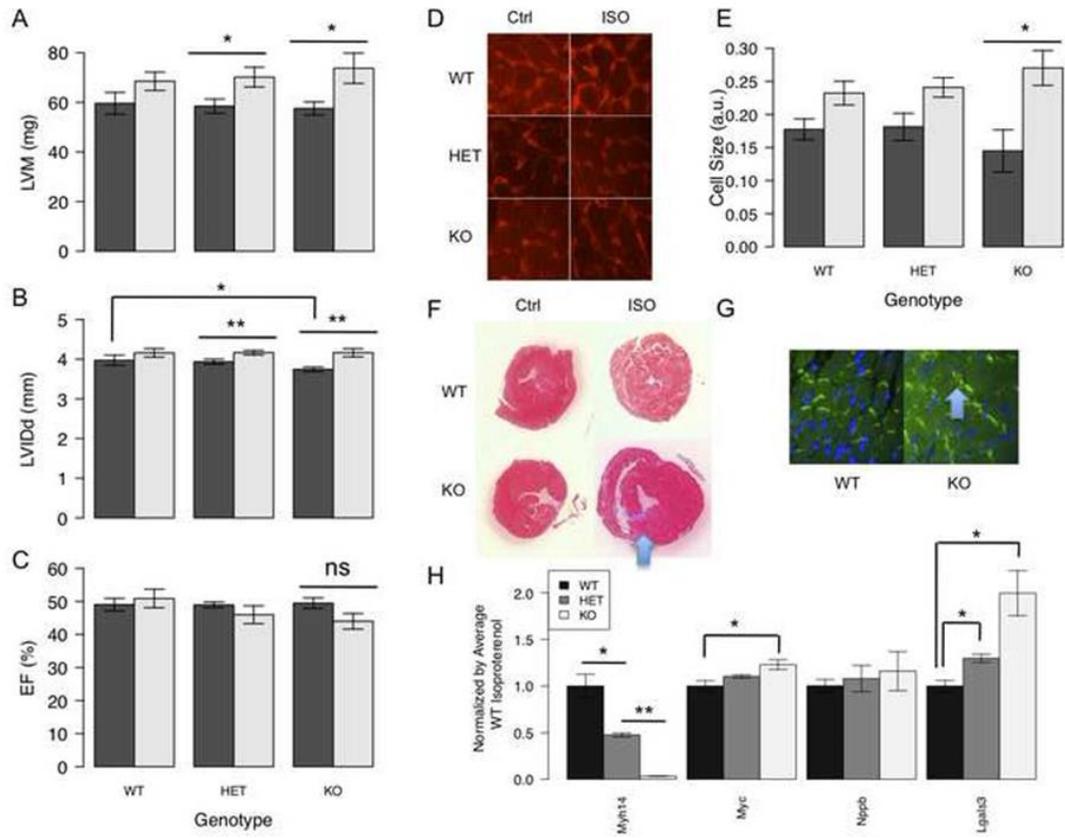


Figure 7.



**Figure 8.**



**Figure 9.**

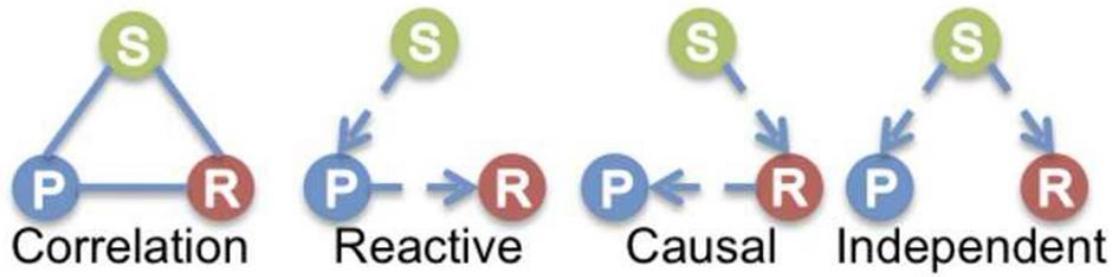


Figure 10.

## References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. *Circulation*. 2015 Jan 27;131(4):e29-322.
2. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KK, et al. Long-term trends in the incidence of and survival with heart failure. *N Engl J Med*. 2002 Oct 31;347(18):1397-402.
3. Lympelopoulos A, Rengo G, Koch WJ. Adrenergic nervous system in heart failure: pathophysiology and therapy. *Circ Res*. 2013 Aug 30;113(6):739-53.
4. Cohn J, Ferrari R, Sharpe N. Cardiac remodeling--concepts and clinical implications: a consensus paper from an international forum on cardiac remodeling. Behalf of an International Forum on Cardiac Remodeling. *Journal of the American College of Cardiology*. 2000;35(8cf6064e-e503-f2d1-9688-594e631a1a17):569-651.
5. Koitabashi N, Kass D. Reverse remodeling in heart failure--mechanisms and therapeutic opportunities. *Nature reviews Cardiology*. 2012;9(2ac7d959-fbf1-f099-cb0f-59305295f364):147-204.
6. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008 Nov 7;322(5903):881-8.
7. Vasani RS, Glazer NL, Felix JF, Lieb W, Wild PS, Felix SB, et al. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. *JAMA*. 2009 Jul 8;302(2):168-78.
8. Smith NL, Felix JF, Morrison AC, Demissie S, Glazer NL, Loehr LR, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and

African ancestry: a prospective meta-analysis from the cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet*. 2010 Jun;3(3):256-66.

9. Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J*. 2011 May;32(9):1065-76.

10. Fox ER, Musani SK, Barbalic M, Lin H, Yu B, Ogunyankin KO, et al. Genome-wide association study of cardiac structure and systolic function in African Americans: the Candidate Gene Association Resource (CARE) study. *Circ Cardiovasc Genet*. 2013 Feb;6(1):37-46.

11. Suzuki M, Carlson KM, Marchuk DA, Rockman HA. Genetic modifier loci affecting survival and cardiac function in murine dilated cardiomyopathy. *Circulation*. 2002 Apr 16;105(15):1824-9.

12. Wheeler FC, Fernandez L, Carlson KM, Wolf MJ, Rockman HA, Marchuk DA. QTL mapping in a mouse model of cardiomyopathy reveals an ancestral modifier allele affecting heart function and survival. *Mamm Genome*. 2005 Jun;16(6):414-23.

13. Chu PL, Keum S, Marchuk DA. A novel genetic locus modulates infarct volume independently of the extent of collateral circulation. *Physiol Genomics*. 2013 Sep 3;45(17):751-63.

14. Wang S, Zhang H, Wiltshire T, Sealock R, Faber JE. Genetic dissection of the *Canq1* locus governing variation in extent of the collateral circulation. *PLoS One*. 2012;7(3):e31910.

15. McNally EM, Barefield DY, Puckelwartz MJ. The genetic landscape of cardiomyopathy and its role in heart failure. *Cell Metab*. 2015 Feb 3;21(2):174-82.

16. Bennett BJ, Farber CR, Orozco L, Kang HM, Ghazalpour A, Siemers N, et al. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res.* 2010 Feb;20(2):281-90.
17. Ghazalpour A, Rau CD, Farber CR, Bennett BJ, Orozco LD, van Nas A, et al. Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm Genome.* 2012 Oct;23(9-10):680-92.
18. Parks BW, Nam E, Org E, Kostem E, Norheim F, Hui ST, et al. Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab.* 2013 Jan 8;17(1):141-52.
19. Farber CR, Bennett BJ, Orozco L, Zou W, Lira A, Kostem E, et al. Mouse genome-wide association and systems genetics identify *Asx12* as a regulator of bone mineral density and osteoclastogenesis. *PLoS Genet.* 2011 Apr;7(4):e1002038.
20. Park CC, Gale GD, de Jong S, Ghazalpour A, Bennett BJ, Farber CR, et al. Gene networks associated with conditional fear in mice identified using a systems genetics approach. *BMC Syst Biol.* 2011;5:43.
21. Rau CD, Wang J, Avetisyan R, Romay M, Martin L, Ren S, et al. Mapping Genetic Contributions to Cardiac Pathology Induced by Beta-Adrenergic Stimulation in Mice. *Circ Cardiovasc Genet.* 2014 Dec 5.
22. Liu J, Rigel DF. Echocardiographic examination in rats and mice. *Methods Mol Biol.* 2009;573:139-55.
23. Moran CM, Thomson AJ, Rog-Zielinska E, Gray GA. High-resolution echocardiography in the assessment of cardiac physiology and disease in preclinical models. *Exp Physiol.* 2013 Mar;98(3):629-44.

24. Galindo CL, Skinner MA, Errami M, Olson LD, Watson DA, Li J, et al. Transcriptional profile of isoproterenol-induced cardiomyopathy and comparison to exercise-induced cardiac hypertrophy and human cardiac failure. *BMC Physiol.* 2009;9:23.
25. Molojavyi A, Lindecke A, Raupach A, Moellendorf S, Kohrer K, Godecke A. Myoglobin-deficient mice activate a distinct cardiac gene expression program in response to isoproterenol-induced hypertrophy. *Physiol Genomics.* 2010 Apr 1;41(2):137-45.
26. Wu J, Bu L, Gong H, Jiang G, Li L, Ma H, et al. Effects of heart rate and anesthetic timing on high-resolution echocardiographic assessment under isoflurane anesthesia in mice. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine.* 2010;29(12):1771-8.
27. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, editor. *Bioinformatics and computational biology solutions using R and Bioconductor.* New York: Springer Science+Business Media; 2005. p. xix, 473 p.
28. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
29. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol.* 2004;3:Article3.
30. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44-57.
31. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009 Jan;37(1):1-13.

32. Yang H, Ding Y, Hutchins LN, Szatkiewicz J, Bell TA, Paigen BJ, et al. A customized and versatile high-density genotyping array for the mouse. *Nat Methods*. 2009 Sep;6(9):663-6.
33. Rau CD, Parks B, Wang Y, Eskin E, Simecek P, Churchill GA, et al. High-Density Genotypes of Inbred Mouse Strains: Improved Power and Precision of Association Mapping. *G3 (Bethesda)*. 2015;5(10):2021-6.
34. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007 Sep;81(3):559-75.
35. Kruijer W, Boer MP, Malosetti M, Flood PJ, Engel B, Kooke R, et al. Marker-based estimation of heritability in immortal populations. *Genetics*. 2015 Feb;199(2):379-98.
36. Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D. FaST linear mixed models for genome-wide association studies. *Nat Methods*. 2011;8(10):833-5.
37. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010 Sep 15;26(18):2336-7.
38. Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. 2011 Sep 15;477(7364):289-94.
39. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073-81.
40. Ma X, Jana SS, Conti MA, Kawamoto S, Claycomb WC, Adelstein RS. Ablation of nonmuscle myosin II-B and II-C reveals a role for nonmuscle myosin II in cardiac myocyte karyokinesis. *Mol Biol Cell*. 2010 Nov 15;21(22):3952-62.

41. Najafi A, Sequeira V, Kuster DW, van der Velden J. beta-adrenergic receptor signalling and its functional consequences in the diseased heart. *Eur J Clin Invest*. 2016 Feb 4.
42. Lok DJ, Lok SI, Bruggink-Andre de la Porte PW, Badings E, Lipsic E, van Wijngaarden J, et al. Galectin-3 is an independent marker for ventricular remodeling and mortality in patients with chronic heart failure. *Clin Res Cardiol*. 2013 Feb;102(2):103-10.
43. Franceschini N, Fox E, Zhang Z, Edwards TL, Nalls MA, Sung YJ, et al. Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. *Am J Hum Genet*. 2013 Sep 5;93(3):545-54.
44. Aouizerat BE, Vittinghoff E, Musone SL, Pawlikowska L, Kwok PY, Olgin JE, et al. GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease. *BMC Cardiovasc Disord*. 2011;11:29.
45. Yoshida T, Yamashita M, Horimai C, Hayashi M. Kruppel-like factor 4 protein regulates isoproterenol-induced cardiac hypertrophy by modulating myocardin expression and activity. *J Biol Chem*. 2014 Sep 19;289(38):26107-18.
46. Rau CD, Lusic AJ, Wang Y. Genetics of common forms of heart failure: challenges and potential solutions. *Curr Opin Cardiol*. 2015 May;30(3):222-7.
47. Post WS, Larson MG, Myers RH, Galderisi M, Levy D. Heritability of left ventricular mass: the Framingham Heart Study. *Hypertension*. 1997 Nov;30(5):1025-8.
48. Sharma P, Middelberg RP, Andrew T, Johnson MR, Christley H, Brown MJ. Heritability of left ventricular mass in a large cohort of twins. *J Hypertens*. 2006 Feb;24(2):321-4.
49. Wong SC, Sleeper LA, Monrad ES, Menegus MA, Palazzo A, Dzavik V, et al. Absence of gender differences in clinical outcomes in patients with cardiogenic shock complicating acute

myocardial infarction. A report from the SHOCK Trial Registry. *J Am Coll Cardiol*. 2001 Nov 1;38(5):1395-401.

50. Vinhas M, Araujo AC, Ribeiro S, Rosario LB, Belo JA. Transthoracic echocardiography reference values in juvenile and adult 129/Sv mice. *Cardiovasc Ultrasound*. 2013;11:12.

51. Anderson M, Moore D, Larson D. Comparison of isoproterenol and dobutamine in the induction of cardiac hypertrophy and fibrosis. *Perfusion*. 2008 Jul;23(4):231-5.

52. Constantinides C, Mean R, Janssen BJ. Effects of isoflurane anesthesia on the cardiovascular function of the C57BL/6 mouse. *ILAR J*. 2011;52(3):e21-31.

53. Wu J, Bu L, Gong H, Jiang G, Li L, Ma H, et al. Effects of heart rate and anesthetic timing on high-resolution echocardiographic assessment under isoflurane anesthesia in mice. *J Ultrasound Med*. 2010 Dec;29(12):1771-8.

54. Francis GS, Desai MY. Contractile reserve: are we beginning to understand it? *JACC Cardiovasc Imaging*. 2008 Nov;1(6):727-8.

55. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006 Aug 25;126(4):663-76.

56. Liao X, Haldar SM, Lu Y, Jeyaraj D, Paruchuri K, Nahori M, et al. Kruppel-like factor 4 regulates pressure-induced cardiac hypertrophy. *J Mol Cell Cardiol*. 2010 Aug;49(2):334-8.

57. Leenders JJ, Pinto YM, Creemers EE. Tapping the brake on cardiac growth-endogenous repressors of hypertrophic signaling. *J Mol Cell Cardiol*. 2011 Aug;51(2):156-67.

58. van Nas A, Pan C, Ingram-Drake LA, Ghazalpour A, Drake TA, Sobel EM, et al. The systems genetics resource: a web application to mine global data for complex disease traits. *Front Genet*. 2013;4:84.

59. Yoshida M, Ohkusa T, Nakashima T, Takanari H, Yano M, Takemura G, et al. Alterations in adhesion junction precede gap junction remodelling during the development of heart failure in cardiomyopathic hamsters. *Cardiovasc Res.* 2011 Oct 1;92(1):95-105.
60. Wang X, Gerdes AM. Chronic pressure overload cardiac hypertrophy and failure in guinea pigs: III. Intercalated disc remodeling. *J Mol Cell Cardiol.* 1999 Feb;31(2):333-43.
61. Hoogeboom D, Essers MA, Polderman PE, Voets E, Smits LM, Burgering BM. Interaction of FOXO with beta-catenin inhibits beta-catenin/T cell factor activity. *J Biol Chem.* 2008 Apr 4;283(14):9224-30.

## Chapter 5

**A systems genetics approach to identify genetic pathways and key drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in mice.**

**Chapter 5 Preface:** This chapter is an insertion of manuscript that is currently in preparation for submission at Cell Systems. The goal of this project was to identify modules, a cluster of genes within the network that are strongly related to each other, that were significantly correlated with heart failure traits to identify novel genes and molecular pathways contributing heart failure. Utilizing gene co-expression network analysis on left ventricular transcriptomes generated from 93 heart failure HMDP strains following exposure to isoproterenol for 21 days to induce cardiac hypertrophy, our group identified module 5 as significantly correlated with the heart failure related traits total heart weight, liver weight, lung weight and cardiac fibrosis. As a collaborator on this project I validated the predicted relationship of *Adamts2* to the module 5 genes *Nppa*, *Tnc*, *Mfap2* and *Kcnv2* using a combination of siRNA knockdown in primary neonatal rat cardiomyocytes and RT-qPCR. I also showed that knockdown of *Adamts2* expression in cardiomyocytes inhibits cardiac hypertrophy *in vitro*. In addition, I have also had a major role in the preparation of the figures and text for the manuscript related to this project.

**A systems genetics approach to identify genetic pathways and key drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in mice**

Authors: Christoph D. Rau<sup>1,3\*</sup>, Milagros C. Romay<sup>1\*</sup>, Jessica J-C. Wang<sup>2</sup>, Marc Santolini<sup>4</sup>, Shuxun Ren<sup>3</sup>, Alain Karma<sup>4</sup>, James N. Weiss<sup>2</sup>, Yibin Wang<sup>3</sup>, Aldons J. Lusis<sup>1,2,5</sup>

<sup>1</sup>Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, CA 90095, USA

<sup>2</sup>Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA

<sup>3</sup>Departments of Anesthesiology, Physiology and Medicine, Cardiovascular Research Laboratories, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

<sup>4</sup>Center for Interdisciplinary Research on Complex Systems, Department of Physics, Northeastern University, Boston, MA, USA

<sup>5</sup>Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

\*Authors contributed equally to this work.

## **Abstract**

Heart Failure (HF) is a complex disease involving numerous environmental and genetic factors. We previously reported a genetic analysis of HF traits in a population of inbred mouse strains treated with isoproterenol to mimic catecholamine-driven cardiac hypertrophy. We now present a systems-level analysis in which we perform co-expression network modeling in left ventricular transcriptomes from these mice. We describe the features of the overall network but focus on a module identified in treated hearts that is strongly related to cardiac hypertrophy and pathological remodeling. Using the causal modeling algorithm NEO, we identified the gene *Adamts2* as a putative regulator of the module and validated its role using siRNA-mediated knockdown in neonatal rat ventricular myocytes. As expected, *Adamts2* silencing regulated the expression of the genes residing within the module and impaired isoproterenol-induced cellular hypertrophy. Our results provide a view of higher order interactions in HF with potential for diagnostic and therapeutic insights.

## **Introduction**

Heart failure (HF) is characterized by impaired heart function, cardiac hypertrophy and chamber remodeling, leading to pulmonary and hepatic congestion (Frangogiannis, 2012). It is a very common disorder with a lifetime risk of 1 in 9 (Mudd and Kass, 2008). Despite significant heritability, genome-wide association studies (GWAS) involving tens of thousands of patients have had only modest success, likely due to the complex, heterogeneous nature of the disease [reviewed in Rau et al., 2015a]. These complexities can be minimized in genetic studies of model organisms such as mice, and classical quantitative trait locus (QTL) linkage analyses in mice have identified a number of novel HF-related genes (McNally et al., 2015; Wheeler et al., 2009).

We have shown that a GWAS approach can be applied to populations of common inbred strains of mice if associations are corrected for population structure (Bennett et al., 2010). We have studied a population of over 100 commercially available inbred strains of mice selected for diversity, constituting a resource which we have termed the Hybrid Mouse Diversity Panel (HMDP). The mapping resolution of this approach is at least an order of magnitude better than traditional QTL analysis involving genetic crosses, and it has now led to the identification of novel genes for a number of traits (Bennett et al., 2010, 2015; Farber et al., 2011; Hui et al., 2015; Orozco et al., 2012; Parks et al., 2013, 2015; Rau et al., 2015b) . We recently applied this approach to identify loci and genes which contribute to HF traits in an isoproterenol (ISO) model. Isoproterenol, which models the chronic  $\beta$ -adrenergic stimulation that often occurs in human HF, was administered by an implanted pump for 3 weeks. Association analysis identified both known and novel genes contributing to hypertrophy and cardiac fibrosis (Rau et al., 2015b).

We now report an analysis in the HMDP Heart Failure study in which we seek to understand genes and pathways which contribute to HF through the modeling of biological networks. We carried out a global transcriptome analysis of the HMDP population before and after treatment with ISO and used an improved version of our previously described co-expression network analysis method, Maximal Information Component Analysis (MICA) (Rau et al., 2013), to form modules of functionally-related genes based on the transcriptome data.

Several modules which showed significant association to HF-related phenotypes were identified. We focused our analysis on a module based on treated expression data since it exhibited striking correlations with a number of heart failure traits, including ventricular hypertrophy and chamber remodeling. Moreover, it contained a number of genes previously implicated in HF, such as *Nppa* and *Timp1*. We then applied the NEO algorithm to develop a directed network with predicted causal interactions among the module genes (Aten et al., 2008). The results suggested that *Adamts2*, a metalloproteinase not previously associated with HF, plays a key role in driving the reaction of other genes in the module in response to ISO stimulation. Using an *in vitro* model, we validated several of these causal links and demonstrated that *Adamts2* expression affected several proxy measurements of cardiac hypertrophy, suggesting that it is a novel regulator of HF and a key member of its module.

## **Results**

### **Gene Network Analysis Using Weighted Maximal Information Component Analysis**

Previous research (Farber, 2013; Ghazalpour et al., 2012) using the HMDP has benefited from the use of systems-level transcriptomics to generate mRNA co-expression networks. We previously reported an unbiased gene network construction algorithm termed Maximal Information Component Analysis (MICA), that captures both linear and nonlinear interactions within the data and allows genes to be spread proportionally across multiple modules (Rau et al., 2013). This approach allows ready identification of module hubs and other important module genes by examining each genes membership within a given module (termed Module Membership, or MM). Past research (Langfelder and Horvath, 2008; Zhang and Horvath, 2005) has demonstrated that weighted network construction algorithms, in which all edges are included in the analysis, have greater versatility and power than unweighted algorithms, in which edges are included or excluded based on a hard threshold, and we report here (see Methods) a modified weighted form of MICA, which we term wMICA. We describe here the application of wMICA to the analysis of heart failure using gene expression data across 92 inbred strains of mice from the HMDP Heart Failure study.

Left ventricular tissue from the HMDP strains (Supplementary Table 1) were processed using Illumina Mouse Ref 2.0 gene expression arrays. In order to focus our efforts on genes which were likely to be members of a module, probes were filtered to transcripts which were significantly expressed in at least 25% of samples and had a coefficient of variation of at least 5%. This resulted in a final set of 8,126 probes, representing 31.6% of the total probes on array. Three gene networks with 20 modules each were generated from these data: one based only on

transcripts from the untreated hearts, one based only on the treated hearts, and a third based on the change in gene expression between these two conditions.

Two measures were used for the preliminary analyses of these networks. We calculated significant gene ontology (GO) enrichments within each of these modules at several module membership cutoffs using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis et al., 2003; Huang et al., 2009). Significant enrichment for one or more GO terms suggests that the module represents a collection of genes which are biologically related to one another and are less likely to be an artifact of the module identification process. We also used principle component analysis to identify the first principle component [often called the “eigengene”(Langfelder and Horvath, 2008; Rau et al., 2013; Weiss et al., 2012)] of each module. This eigengene can be correlated to HF-related phenotypes to identify modules which, as a whole, are most likely related to specific features of cardiac pathogenesis. As wMICA allows genes to reside within multiple modules and, theoretically, influence the function of every detected module, we used a weighted PCA algorithm to calculate the first *weighted* principle component of each module based on each gene's module membership with that module. The weighted eigengene was then correlated to a number of HF-related phenotypes from the HMDP panel. They include 7 organ weights, 8 echocardiographic parameters and 5 plasma traits which, in conjunction with either an organ or functional parameters suggest a change in metabolism status associated with HF.

### **Heart Failure Co-expression Networks**

*MICA modules generated from untreated left ventricular tissue.*

Expression data for a number of genes from the untreated hearts showed strong associations with HF traits. For example, four genes (*Pde4dip*, *Spp1*, *Ppp1r1b* and *Ctstl*) showed significant correlations with RV heart weight and have been previously linked to cardiomyopathies in the literature (Barp et al., 2015; Sun et al., 2013; Uys et al., 2011; Wittköpper et al., 2010) (data not shown). Therefore, the untreated network was first examined to identify possible modules which predispose HF-related pathologies at basal state. Of these, three modules (8,11,13) were observed to have very strong (DAVID score >7) enrichments, while five other modules showed significant (score > 3) enrichments (Supplementary Table 2). Only three modules (7,8,20) had significant ( $P < 5E-4$ ) correlations with HF phenotypes (total and unesterified plasma cholesterol), but only a single module (4) had even suggestive ( $P < 0.01$ ) correlations to an organ weight (right ventricle  $P = 0.004$ ) or echocardiographic phenotype (intraventricular septum  $P = 7E-4$ ) (Supplementary Figure 1). Therefore, the co-expression gene modules identified in the pre-treated hearts appear to have limited correlation with susceptibility to heart failure.

*MICA modules generated from treated left ventricular tissue.*

Thirteen of the modules of the MICA network constructed from treated hearts had significant DAVID enrichments, including three modules (3,4,5) with enrichment scores greater than 10 (Figure 1, Table 1). Three modules (1,5,19) contained at least one significant ( $P < 5E-4$ ) correlation between the eigengene and either an organ weight or echocardiographic parameters (Figure 2, Supplementary Figure 2). Four additional modules (4,12,15,18) contained at least two suggestive ( $P < 0.01$ ) correlations between a module and a HF-related trait. We examined the modules which showed significant correlations to organ or echocardiographic phenotypes in greater detail.

Module 1 was strongly correlated ( $P=5E-4$ ) with right ventricular weight and showed enrichment for cell morphology ( $P=0.007$ ). One hundred forty-four genes have maximal module membership (MM) in module 1, although only 21 genes have over 50% MM and none over 70%. Five of the core genes with over 50% MM showed strong correlation to RV weight (Supplementary Table 3), with two of these genes being previously identified as playing a role in heart development [*Tcf21* (Vicente-Steijn et al., 2015)] or cell expansion [*Mobk1b* (Xiong et al., 2016)]. Only five genes were connected to other genes within the core set of 21 probes with a maximal information criterion (MIC) score of greater than 0.5 (Supplementary Figure 4). This weak internal module structure, with few strong links between genes, made causal analyses difficult within this module.

Module 19, consisting of 65 genes with maximal membership to the module, showed strong correlations with 8 of the 20 HF traits: liver weight, plasma cholesterol, unesterified cholesterol and free fatty acids, fractional shortening, ejection fraction, mean circumferential fiber shortening velocity and mean normalized systolic ejection rate. This suggests a relationship between this module and cardiac contractility and metabolism. Using DAVID, module 19 showed enrichment for the GO terms contractile fibers ( $P=0.008$ ) and vasoconstriction ( $P=0.015$ ). Of the 19 core module 19 genes, *Prkaa2*, one isoform of the active subunit of AMPK is noteworthy as it was identified by a GWAS study from the same HMDP panel as a candidate gene for heart weight (Rau et al., 2015b).

Module 5, with a total of 309 genes exhibiting maximal module membership within the module (Fig. 1), showed strong correlations to 7 of 20 HF traits measured: total heart weight, left and right ventricle weights, left ventricular internal dimension (LVID), lung and liver weight and plasma free fatty acids, suggesting a strong relationship between this module and cardiac

hypertrophy and heart failure (Figure 2). Module 5 showed highly significant GO enrichments for several biological processes: extracellular matrix ( $p < 10^{-18}$ ), secreted signaling ( $p < 10^{-16}$ ) and cell adhesion ( $p < 10^{-9}$ ). Of the 47 probes (42 genes) which possess greater than 70% module membership (Figure 3), twenty genes (47.7%) have previously been described as involved in heart hypertrophy or cardiac remodeling based on transgenic studies in animal models or mendelian forms of HF. These genes include *Nppa*, a well-known marker of cardiac hypertrophy, *Timp1*, an extracellular matrix regulator (Barton et al., 2003; Ikonomidis et al., 2005), and the collagen genes *Coll2a1*, *Coll4a1* and *Col6a2*. We further observed three genes (*Pcolce*, *Sprx*, *Sprx2*) which we had previously identified via GWAS as candidate genes for cardiac phenotype in this panel of mice (Rau et al., 2015b). These interesting features led us to further characterize module 5 in greater detail.

*MICA modules generated from the change in gene expression between control and treated mice (Delta Network).*

We observed the strongest associations between gene modules and GO terms in the MICA network generated from the changes in gene expression induced by isoproterenol and subsequent remodeling, with eleven modules showing significant enrichments (Supplementary Table 4). Four of these modules had DAVID scores greater than 10 and two (6 and 14) had scores greater than 15. Despite strong enrichments for GO terms, correlation of these module's eigengenes with the changes in HF-related phenotypes (Supplementary Figure 3) were only significant or suggestive for changes in unesterified cholesterol or free fatty acids. Notably, no eigengene/trait correlations for any organ weight or echocardiographic parameter had a p-value of less than 0.01.

Although none of the modules from this network showed strong correlations to organ or echocardiographic HF-related phenotypes, we did note that module 14 appeared to significantly overlap with module 5 from the treatment-specific network. Both modules shared the same primary GO enrichment (Secreted Signaling: 18.6 for module 14, as compared to 17.8 for module 5), and a significant (50%,  $P=5.95E-55$ ) overlap between the 'core' genes of each module with module membership greater than 50%, including *Nppa*, *Sprx2* and *Timp1*. Despite these overlaps, however, the eigengene of module 14 did not significantly correlate with any organ or echocardiographic parameter, while the eigengene of module 5 was significantly correlated with many parameters. This suggests that, despite the shared genes between the two modules, the genes which are driving the modules to influence phenotype are different between these two modules.

### **Causality Modeling of Modules Using the Near Edge Orientation Algorithm**

Most co-expression network analysis methods, including wMICA, return a non-directed network since gene-gene relationships are based on association rather than cause/effect modeling. It is possible to convert these non-directed edges into causal links through the addition of additional genetic information. For example, if two genes are highly correlated and one has a strong *cis*-acting expression quantitative trait locus (eQTL), it is consistent with that gene being the driver of the other. The Near Edge Orientation (NEO) (Aten et al., 2008) algorithm uses SNPs as anchors to infer directionality between SNP/Gene/Gene triads based on several possible models (Figure 4A). We applied the NEO algorithm to the largest connected components of the core genes of modules 5 and 19 from the treated MICA network and module 14 from the delta network.

*Abhd1*, a poorly characterized hydrolase, was the predicted driver of the largest connected component of module 19 (Supplementary Figure 5), with its expression driving the expression of *Prkaa2*, *Clec16a* and *Vrk1*. Little or no directionality was found between any of the other genes in the module. In the case of module 5 of the ISO-treated network, 44 of the 47 probes in the core network had at least one edge with significant directionality (Figure 4B). Genes with directed edges were classified into three categories (reactive, intermediary, driver) based upon the number of directed edges traveling from the gene of interest to other genes in the module and the number of directed edges traveling from other genes in the module to the gene of interest (Supplementary Table 5). We observe 12 "reactive" genes in which 75% of their directed edges travel to the gene from other genes in the module. The most reactive gene is *Timp1*, an important marker of left ventricular remodeling and heart failure (Barton et al., 2003; Milting et al., 2008), which is predicted to be affected by 18 other genes in the module. Similarly, we classified 12 genes as drivers based on the observation that at minimum 75% of their directed edges originate from the gene and travel to other genes in the module. Of the 12 drivers, 3 are notable. The first, *Pcolce*, a previously reported GWAS candidate gene (Rau et al., 2015b), has 11 outputs representing 26% of the genes in the core of module 5. The second, *Nox4*, which has a single output to *Mfap5*, has been previously established to be a key contributor to oxidative stress during pressure overload (Kuroda et al., 2010). The third driver, the metalloproteinase *Adamts2*, has not been previously implicated in heart failure or muscle development and has the largest number of directed edges within the module, with 23 causal and 2 reactive (to the exocytosis regulator *Rab15* and the extracellular matrix organizer *Ccdc80*) edges. Furthermore, *Adamts2* showed significant correlations to lung weight (R=0.46, P=4.6E-6), heart weight (R=0.43, P=1.5E-5) and free fatty acids (R=0.43, P=2.0E-5) after ISO stimulation. These results

suggest that *Adamts2* is a novel regulator of cardiac pathology. This prediction has now been validated by functional studies in cardiomyocytes as shown below.

Module 14 from the delta network shared many features with module 5, including a large proportion of shared genes with module membership greater than 50% and similar GO enrichments according to DAVID. Module 5, however, shows strong correlations to multiple phenotypes while module 14 does not. NEO revealed that despite their similarities, the ways in which the two modules are wired are very different (Supplementary Figure S6 vs Figure 4B). Module 5 shows a clear directionality to the module, with one or two major hub drivers positioned upstream of nearly every other gene, including the other driver genes. Module 14 lacks this underlying structure, and in fact, the major drivers of module 5, *Pcolce* and *Adamts2*, are not present in module 14, replaced instead by the exocytosis regulator *Rab31* and the cardiac growth regulator *Coll4a1* (Tao et al., 2012). Similarly, the most reactive gene in module 5, *Timp1*, is now only regulated by two other genes.

### **Validation of the Driver Role of *Adamts2* in Module 5 Using Cultured Cardiomyocytes**

To confirm the NEO predicted regulation of gene expression by *Adamts2*, we measured the expression of a set of 7 predicted downstream targets residing in module 5. These seven genes (*Coll2a1*, *Kcnv2*, *Mfap2*, *Nppa*, *Pcolce*, *Timp1* and *Tnc*) were selected based on three criteria: strength of predicted relationship to *Adamts2*, number of predicted directed edges in module 5, and previous associations with cardiac function and/or cardiovascular disease (Supplementary Table 6). Using siRNA to inactivate *Adamts2*, we achieved a 28-45% decrease in *Adamts2* expression when compared to transfection control (Figure 5A). Following knockdown of *Adamts2*, we measured the expression of the 7 predicted downstream genes and observed significant changes in 4 (*Kcnv2*, *Mfap2*, *Nppa* and *Tnc*) (Figure 5). On average, we

observed 1.4 fold to 1.9 fold induction in *Kcnv2*, 1.3 to 1.7 fold induction in *Tnc* and 2 to 3.8 induction in *Mfap2* following knockdown of *Adamts2* in cardiomyocytes. We also observed approximately an 80% decrease in the expression *Nppa*. These findings suggest that *Adamts2* acts to regulate the expression of *Kcnv2*, *Mfap2*, *Nppa* and *Tnc* under ISO treated conditions in cardiomyocytes. The failure to observe the predicted resources for the other genes is likely due either to insufficient knockdown of *Adamts2* or incorrect prediction by the NEO analysis.

***Adamts2* acts as a regulator of  $\beta$ -adrenergic induced cardiac hypertrophy *in vitro*.**

As module 5 is significantly correlated with changes in numerous HF related traits, we hypothesized that changes in the expression in *Adamts2*, a predicted driver of module 5, would alter cardiomyocyte size and/or viability in response to ISO. Indeed, following treatment with ISO those cells transfected with the control siRNA nearly doubled in cell area (Figure 6A and 6B). When compared to transfected control cells, cells that expressed less *Adamts2* due to siRNA transfection were smaller following treatment with ISO, being about the same size as the control cells without ISO stimulation (Figure 6A and 6B).

At the molecular level, treatment with ISO induced the expression of the hypertrophic markers *Nppa* and *Nppb*, which rose 2.4 fold and 5.7 fold, respectively, in cells transfected with the control siRNA. Knockdown of *Adamts2* strongly impaired the induction of these genes, as *Nppa* induction was reduced approximately by 75% and *Nppb* expression was reduced approximately 65% (Figure 6C and 6D). The significant response of the hypertrophic markers to relatively modest changes in *Adamts2* expression under an adrenergic stimulus suggests a non-linear relationship between the genes. As such these findings indicate that *Adamts2* acts as a novel regulator of  $\beta$ - adrenergic induced cardiac hypertrophy in cardiomyocytes.

## Discussion

We report network modeling of the molecular pathways contributing to HF in an ISO-treated mouse model. A modified, weighted form of a previously reported network algorithm, MICA, was used to construct co-expression networks from the transcriptomes of left ventricular samples taken from 92 highly diverse and well-characterized inbred strains of mice after isoproterenol treatment to mimic the chronic catecholamine activation in heart failure. We also constructed a network based on 92 matched control strains of mice as well as a network based on the change in gene expression in these 92 strains after ISO stimulation. Three modules from these networks were identified as being of particular interest. Applying the NEO algorithm, we identified drivers for each of these modules, including the gene *Adamts2* which is predicted to directly regulate 45% of the genes residing within its module. siRNA knockdown of *Adamts2* in neonatal rat ventricular cardiomyocytes confirmed the regulation of the module genes *Nppa*, *Kcnv2*, *Mfap2* and *Tnc* by *Adamts2* and revealed that *Adamts2* acts as a regulator of cardiac hypertrophy.

Prior efforts to analyze the transcriptome networks underlying heart failure and related cardiomyopathies in human studies have met with limited successes. The difficulty in obtaining large numbers of high quality human samples has limited such studies (Camargo and Azuaje, 2007; Moreno-Moral et al., 2013) to small-scale which are underpowered to identify highly relevant modules with strong GO enrichments and correlations to phenotypes of interest. Data interpretation can be further complicated by significant complexity in etiology and disease progression, resulting in high degree of heterogeneity even when a large number of samples were included in a study (Dewey et al., 2011; Drozdov et al., 2013). Additionally, consistent phenotyping protocols are harder to maintain over multiple centers, personnel and years of data

collection. Finally, human studies are typically only able to obtain tissue samples from hearts with extremely late-stage HF, which masks the pathways involved in the initial stage in favor of the reactive pathways which result after months to years of HF progression.

The use of the HMDP to study HF using transcriptome network modeling minimizes many of these issues. The HMDP is a mouse resource which has previously been successfully used in network biology (Calabrese et al., 2012; Park et al., 2011; Rau et al., 2013) and which has sufficient power to perform GWAS analyses as well (Ghazalpour et al., 2012; Parks et al., 2013; Rau et al., 2015b). We reduced environmental heterogeneity by raising the 776 mice used in the HF HMDP study under the same conditions and treating with the same stressor. Experimental variability was controlled by using the same operator and machine for each animal when gathering phenotypic traits.

Although we constructed MICA networks from three separate sets of gene expression data, we found that the treated network returned the most relevant modules. The control network had generally weaker GO enrichments (Table S2) and only a single module with even suggestive ( $P < 0.01$ ) correlations with organ weights or echocardiographic parameters. Thus, although individual genes exhibited significant correlations with basal HF-related traits, basal genetic variation does not appear to be a reliable predictor of gene modules which correlate to untreated phenotypes. The delta network had GO enrichments on par with the treated modules (Table S4), but, no significant correlations with the HF phenotypes. Hierarchical clustering analysis (Figure S7) demonstrated that 11 of the 20 modules from the delta network are highly correlated ( $R > 0.8$ ) to one another. This suggests that the identified modules in the network are in fact largely reflecting a single 'mega-module' of genes which have reacted strongly to ISO stimulation but whose underlying genetic variability has been masked to such a degree that they cannot be

differentiated into phenotypically-relevant modules. Thus, we conclude that the network based on ISO-treated gene expression is most useful for network modeling at least in this context.

For further analyses, we focused on module 5 of the ISO-treated network since it was significantly correlated with the traits total heart weight, lung weight, liver weight, and plasma free fatty acids as well as some functional parameters. In addition to its strong correlation to HF related traits, we noted a strong enrichment for genes previously associated with cardiovascular disorders and development. Out of the 19 genes previously associated with cardiovascular disorders and development 15 (*Col6a2*, *Col14a1*, *Comp*, *Cx3cl1*, *Dkk3*, *Egfr*, *Itga11*, *Nox4*, *Olfml3*, *Pcolce*, *Ptn*, *Srpx2*, *Timp1*, *Tnc* and *Tubb2b*) have been implicated in the development of heart failure and heart failure related traits. Expression of 6 genes (*Col14a1*, *Dkk3*, *Olfml3*, *Ptn*, *Srpx2* and *Timp1*) of these 13 genes is predicted to be cardioprotective as loss of their expression promotes cardiac remodeling [*Col14a1* (Tao et al., 2012), *Dkk3* (Bao et al., 2015; Zhang et al., 2014), and *Timp1* (Ikonomidis et al., 2005)] and decreases angiogenesis and neovascularization [*Olfml3* (Miljkovic-Licina et al., 2012), *Srpx2* (Miljkovic-Licina et al., 2009), and *Ptn* (Li et al., 2007)]. In contrast, expression of 6 genes (*Col6a2*, *Cx3cl1*, *Egfr*, *Itga11*, *Nox4*, and *Tnc*) are implicated as maladaptive as their overexpression leads to increase in cardiac remodeling [*Col6a2* (Grossman et al., 2011), *Cx3cl1* (Gu et al., 2015; Xuan et al., 2011), *Egfr* (Messaoudi et al., 2012), *Pcolce* (Kessler-Icekson et al., 2006), and *Tnc* (Hessel et al., 2009; Nishioka et al., 2010)] and increased oxidative stress (*Nox4*) and fibrosis (*Itga11*) (Talior-Volodarsky et al., 2012) during heart failure. The remaining 4 genes have been implicated in the development of vascular disorders such as generalized arterial calcification of infancy [*Enpp1* (Hofmann Bowman and McNally, 2012)] and thoracic abdominal aneurysm [*Mfap5* (Barbier et al., 2014)]. Two genes (*Fbln1* and *Snai1*) have been implicated in cardiovascular development through

regulation of ventricular morphogenesis [*Fbln1* (Cooley et al., 2012)] and cardiac valve formation [*Snai1* (Tao et al., 2011)].

In addition to its enrichment for known HF genes, module 5 contained a second intriguing characteristic. Despite the fact that ISO activates the neurohormonal signaling cascade, many of the genes residing within module 5 with the exception of *Nppa*, are not directly related to the canonical  $\beta$ -adrenergic signaling. Rather, upon further inspection, we find that approximately one-third (29%) have previously been implicated in inflammatory signaling, integrin signaling or TGF- $\beta$  signaling. Four genes (*Cx3cl1*, *Capn5*, *Cercam* and *Ccdc80*) are related to activation of inflammatory signaling. In particular an inhibitor of *Cx3cl1*, known as fractalkine, is well established to suppress progression of HF in both MI and TAC models of HF in mice (Xuan et al., 2011). This observation provides potential mechanistic cross-talk between bAR signaling and inflammatory pathway in the progression of heart failure. Five module 5 genes (*Comp*, *Fbln1*, *Itgal1*, *Svepl* and *Tnc*) have either direct interactions with integrins expressed on the cardiomyocyte surface or are integrin proteins. In addition to direct interaction with integrins, several module 5 genes interact with proteins known contribute to integrin signaling, including Fibrillin 1 [*Mfap2* and *Mfap5* (Hanssen et al., 2004)], Decorin [*Col14a1* (Ehnis et al., 1997)], Fibronectin [*Col6a2* (Tillet et al., 1994)] and *FAK* [*Srpx2* (Tanaka et al., 2009)]. All these findings suggest integrin mediated signaling plays a potential role in ISO induced stress response and subsequent pathology in heart.

Finally, we identified 13 genes representing 3 classes of players in TGF- $\beta$  signaling cascade at multiple points from systems analysis. Transforming growth factor –  $\beta$  (TGF- $\beta$ ) signaling plays an important role in the pathogenesis of cardiac remodeling through its effects on cardiomyocytes, immune and mesenchymal cells (Dobaczewski et al., 2011). In the first

category, we find genes which activate or enhance the TGF- $\beta$  signaling cascade. These include, *Itgbl1* (Li et al., 2015), *Itga11* (Talior-Volodarsky et al., 2012), *Olfml3* (Miljkovic-Licina et al., 2012) and *Pcoale* (Moali et al., 2005). The second class of genes are those that inhibit or antagonize the cascade such as *Nbl1* (Hung et al., 2012) and *Dkk3*. The third class of genes are potential downstream targets of TGF- $\beta$  signaling and include *Coll4a1* (Arai et al., 2002), *Enpp1* (Goding et al., 2003), *Fbln1* (Chen et al., 2013), *Nox4* (Yan et al., 2014), *Ptn* (Kosla et al., 2013), *Timp1* (Hall et al., 2003), and *Tnc* (Jinnin et al., 2004). These data implies that ISO induced cardiac pathology involves multiple players up-and down-stream of TGF- $\beta$  signaling.

Using NEO causal modeling, we identified *Adamts2* as a key driver of module 5. *Adamts2* is a member of the “disintegrin and metalloproteinase with thrombospondin motifs” family of metalloproteases (Tortorella et al., 2009). Mutations in *Adamts2* can cause the disorder Ehler Danlos Type VIIC (dermatosparaxis type) which is characterized by severe skin fragility and joint laxity (hyperextensibility) (Porter et al., 2005). Using siRNA knockdown of *Adamts2* in cardiomyocytes we identified 5 new targets of *Adamts2*; *Kcnv2*, *Mfap2*, *Tnc*, *Nppa* and *Nppb*. Of these, *Tnc*, *Nppa* and *Nppb* are well associated with the development of heart failure, while *Kcnv2* and *Mfap2* represent two novel genes not previously implicated in the development of heart failure. Recent work has suggested that *Adamts2* may act as a upstream regulator of TGF- $\beta$  signaling through direct cleavage of the proteins TGF $\beta$  RIII and DKK3 (Bekhouche et al., 2016). *Dkk3*, a known cardioprotective molecule in heart failure also resides within module 5, a module enriched for TGF- $\beta$  signaling genes. Using siRNA knockdown of *Adamts2* in NRVMs, we show that expression of *Adamts2* is crucial to ISO induced cardiac hypertrophy *in vitro* which may in part be due to its regulation of *Dkk3* cleavage and expression of the novel module 5 target genes.

In summary, combining genetic variants, transcriptome profiling and well-controlled pathological stress, we are able to reveal cardiac gene networks associated with heart failure, and uncover a previously unknown player cardiac hypertrophy regulation, *Adamts2*. It is likely that the ISO-induced cardiac pathology reflects only some aspects of the complex spectrum of human heart disease and it will be of interest to compare our results with those generated by other stressors such as an  $\alpha$ -adrenergic agonist such as angiotensin or pressure-overload by trans-aortic constriction. It will also be of interest to compare networks generated in animal models with those observed in human studies, although, as discussed above, the latter will likely reflect the end stages of the disease. With its limitations, the datasets and the insights from this study will offer a wealth of information for future explorations.

## **Materials and Methods**

### **Online database**

All results and data can be accessed at <http://systems.genetics.ucla.edu/data>. Microarray data may also be accessed at the Gene Expression Omnibus using accession ID: GSE48760.

### **Mice and isoproterenol treatment**

The following mouse strains were obtained from The Jackson Laboratory and then bred in our colony: 29 common inbred strains (129X1/SvJ, A/J, AKR/J, BALB/cJ, BALB/cByJ, BTBR T+ tf/J, BuB/BnJ, C3H/HeJ, C57BL/6J, C57BLKS/J, C58/J, CBA/J, CE/J, DBA/2J, FVB/NJ, KK/HIJ, LG/J, LP/J, MA/MyJ, NOD/ShiLtJ, NON/ShiLtJ, NZB/BINJ, NZW/LacJ, PL/J, RIIS/J, SEA/GnJ, SJL/J, SM/J, SWR/J) and 69 RI lines [RI (number of strains) - BXD (40), AXB(8), BXA(10), BXH(5), CxB(6)]. All animal experiments were conducted following guidelines established and approved by the University of California, Los Angeles Institutional Animal Care and Use Committee. All mice have been previously genotyped at over 130,000 locations. Isoproterenol (20 mg per kg body weight per day) was administered for 21 d in 8- to 10-week-old female mice using ALZET osmotic minipumps, which were surgically implanted intra-peritoneally.

### **RNA Isolation and Microarray processing**

Following homogenization of left ventricular tissue samples in QIAzol, RNA was extracted using the Qiagen miRNAeasy extraction kit, and verified as having a RIN>7 by Agilent Bioanalyzer. Two RNA samples were pooled for each strain/experimental condition and arrayed on Illumina Mouse Reference 8 version 2.0 chips. Analysis was conducted using the Neqc algorithm included in the limma R package(Smyth, 2005) and batch effects addressed through the use of COMbat (Johnson et al., 2007).

## MICA

We used a modified form of Maximal Information Component Analysis (Rau et al., 2013) to form gene networks on the LV transcriptomes generated from the ISO-treated mice. Probes were limited to the 8126 probes which were both expressed in at least 25% in both the untreated or treated data, and whose coefficient of variation was greater than 5%. The MICA algorithm consists of two parts. For the first step, relationships between genes are determined using the MINE algorithm (Reshef et al., 2011). This step was not altered from Rau 2013, although we used the MINERVA R package (Albanese et al., 2013) instead of the original Java implementation due to MINERVA's significant improvement in run-time. The second step of the MICA algorithm utilizes the ICMg algorithm (Parkkinen and Kaski, 2010). We have modified this algorithm to allow for weighted edges as described below.

ICMg is an iterative process. For each iteration, each edge is independently interrogated utilizing Gibbs sampling with the following equation:

$$p(z_0|\{z\}', \{L\}', \alpha, \beta) \propto \frac{n'_{z_0+\alpha}}{N' + C_\alpha} * \frac{(q'_{z_0i_0} + \beta)(q'_{z_0j_0} + \beta)}{(2n'_{z_0} + 1 + M\beta)(2n'_{z_0} + M\beta)}$$

where  $\{L\}'$  is the set of all links excluding the one being interrogated,  $\{z\}'$  is the set of module assignments for the links excluding the link being interrogated,  $n_z$  is the count of links assigned to component  $z$ ,  $i$  and  $j$  represent the genes linked by edge  $z_0$  and  $q_{z_i}$  counts the module-node co-occurrences between module  $z$  and node  $i$ .  $C$  is the total number of modules, and  $M$  is the total number of nodes.  $\alpha$  and  $\beta$  are control parameters which modify the overall distribution of module sizes and the average module membership per gene per module. In order to allow for weighted edges, we altered how the matrix  $q$  was updated. Previously,  $q_{z_i}$  and  $q_{z_j}$  were incremented by 1 each time an edge was placed into a module in an iteration. Instead, we increase these entries in the  $q$  matrix by the MIC score of the edge ( $L_{ij}$ ), and instead of dividing

$q$  by the number of iterations in order to get the proportional module memberships for each gene within each module, we divided each element of  $q$  by its respective column sum

Other than the modification described above to allow for weighted edges, analysis was performed as described in Rau et al., 2013. Briefly, transcriptomes were processed using the MINE algorithm to determine relationships between each gene in the dataset. All edges with a MIC score of greater than 0.3 were selected and used as the input for the weighted ICMg algorithm. The ICMg algorithm then determined the proportional module membership for each gene. Weighted principle components (“eigengenes”) of each module were determined through the use of the `dudi.pca` function from the `ade4` R package (Chessel et al., 2004) and compared to HF-related phenotypes using the `heatmap` function of WGCNA (Langfelder and Horvath, 2008). For all control variables, the standard values were used.

Choosing the best %MM cutoff involved balancing the desire for small modules (which occur at high %MM cutoffs) and preservation of the eigengenes returned by MICA (which are highest at low %MM cutoffs). We selected several %MM thresholds and selected the highest %MM cutoff for which the correlation between the eigengenes at that cutoff and the weighted eigengenes were over 90%. This cutoff was 70%. The latest version of the `wMICA` algorithm may be found at <https://github.com/ChristophRau/wMICA>.

## **NEO**

The Network Edge Orienting (NEO) R software package (Aten et al., 2008) uses transcriptome and genotype information to infer a causal link between two genes. We applied this analysis to the genes in module 5. Probes with over 70% MM were included in the analysis. For each edge contained within module 5, NEO was performed using those genes as well as any SNP for which either gene had an eQTL with significance less than  $1E-4$ . Standard protocol for

NEO was used, and we kept any SNP/gene/gene combination which yielded a LEO\_NB.AtoB or LEO\_NB.BtoA score greater than 0.75 and for which  $\text{mlogp.M.AtoB}$  or  $\text{mlogp.M.BtoA}$  was less than 0.05.

In order to make a final determination of which direction an edge went, we examined all SNP/gene/gene combinations kept in the above step. For each edge, we examined only combinations which included that edge. If all of the combinations were either 'forward' (gene A affects gene B) or 'reverse' (gene B affects gene A), then the edge was classified as either 'forward' or 'reverse', respectively. Otherwise, the difference between the sum of the LEO\_NB.AtoB scores (for the 'forward' combinations) and the sum of the LEO\_NB.BtoA scores (for the 'reverse' combinations) was determined. If this difference was larger than 1, then the edge was classified as either 'forward' or 'reverse' (depending on which sum was larger). Otherwise, the directionality of the edge could not be determined and it was classified as 'undirected.'

### **Cell Culture and Treatments**

Neonatal Rat Ventricular Cardiomyocytes (NRVM) were isolated as previously described (Brown et al., 2005). Following isolation the myocytes were plated in DMEM containing 10% fetal bovine serum (FBS) containing 1% antibiotics overnight. For the rest of the culture period the cells were maintained in DMEM containing 1% ITS (Fisher - CB-40351) and 1% antibiotics. For some studies cells were treated with either 60uM Isoproterenol (Sigma - I6504) for 48 hours.

*Adamts2* knockdown experiments were performed using IDT DsiRNA (*Adamts2* siRNA -1 # RNC.RNAI.N001137622.12.1, *Adamts2* siRNA-2 #RNC.RNAI.N001137622.12.5, Negative Control - DS NC1) at the indicated concentrations. All transfections were performed with Invitrogen Lipofectamine RNAimax reagent.

### **cDNA synthesis and qRT-PCR**

mRNA reverse transcription was performed using the Applied Biosystems Reverse Transcription Kit. Quantitative PCR was performed using the KAPA SYBRFast Master Mix in a Roche LightCycler 480 instrument.

### **Acknowledgments**

This work was supported by NIH grants HL110667, HL123295, and HL114437. CDR was supported by NIH training grant T32HL69766, JW was supported by NIH training grant HL007895 and MCR was supported by Ruth L. Kirschstein National Research Service Award T32GM718539. We would like to thank Doug Chapski, Mario Deng, Adriana Huertas-Vazquez, Hrayr Karaguesian, Jonathan Hoffman, Jim O'hearn, Tom Vondriska and Nick Wisniewski for fruitful discussions regarding this manuscript.

## Figure Legends

**Figure 1: Visualization and functional enrichment of the left ventricular gene co-expression network.** In the whole network 8126 transcripts were arranged into 20 distinct modules by the network construction algorithm, MICA. For the ease of visualization, only probes with a module membership of greater than 50% in a module (4,392) were included. For further ease of visualization, probes with a module membership of over 50% in module 9 (2,783) were removed as they tended to mask the relationships between the genes in the smaller modules. Circles (nodes) represent transcripts and lines (edges) represent significant (MICA score  $> 0.35$ ) relationships between the two nodes. Probes are arranged by correlation strength so small distances indicate high correlation and node color indicate module membership. The group of transcripts corresponding to each module was analyzed using DAVID to identify any significant enrichment for biological processes. The most significant biological processes for that particular module have been labeled in the color corresponding to the module.

**Figure 2: Module - HF Trait Correlation.** A weighted PCA algorithm was used to calculate the first weighted principle component of each module. These weighted first principle components were then correlated to the HF-related phenotypes. Strength of correlation is indicated by color, while the p-value of the correlation is indicated by the numbers in the table

**Figure 3: Visualization and functional enrichment of Module 5.** In the whole network, 309 probes contain a maximal membership to module 5. For the ease of visualization only those 47 probes (42 genes) which possess a module membership of 70% or greater are shown (defined as the core module). Diamonds indicate those genes which have been previously implicated in heart disease while circles indicate those genes which have not been previously implicated in heart disease. Color of the diamonds and circles indicate GO category membership of that gene

with blue indicating glycoprotein, red signaling, green indicating extracellular region and grey no significant go category membership.

**Figure 4: Near Edge Orientation Algorithm.** **A.** A schematic depicting the way in which the NEO algorithm uses genetic markers (S) as anchors to infer directionality between two genes (A and B). **B.** Module 5 after NEO. 44 probes of the module, representing 93.6% of the core module showed at least one edge with significant inferred directionality. Circles indicate the nodes while arrows indicate the inferred directionality by NEO between two nodes. Color of the nodes indicate the percent of edges with inferred directions to or from the node with green indicating 75% or greater inferred edges originating from the node, yellow indicating roughly equal percentage of edges originating from and traveling to the node and red indicating that 75% or greater of the edges originating from an alternate node.

**Figure 5: *Adamts2* alters the expression of module 5 genes in cardiomyocytes.** NRVMs were transfected with either control siRNA or *Adamts2* siRNA. Following overnight transfection NRVMs were treated with isoproterenol (60 $\mu$ M) containing media for 48 hours. mRNA levels were quantified using RT-qPCR. Plots show average  $\pm$  SEM. N =9

**Figure 6: *Adamts2* regulates cardiac hypertrophy *in vitro*.** **A-D** NRVMs were transfected with either control siRNA or *Adamts2* siRNA. Following overnight transfection NRVMs were treated with either control or isoproterenol (60 $\mu$ M) containing media for 48 hours. **A** Representative images of cardiomyocytes transfected with either control or *Adamts2* siRNA following 48 hour treatment with control or isoproterenol containing media. **B.** Quantification of cardiomyocyte cell size following transfection with either control or *Adamts2* siRNA and a 48 hour treatment with control or isoproterenol containing media. Plots show average  $\pm$  SEM. N

=6. **C and D.** mRNA expression levels of *Nppa* and *Nppb* were quantified using RT-qPCR.

Plots show average +/- SEM. N = 8-9.

**Table 1: Top Gene Ontology Enrichments in Modules of the left ventricular Gene Coexpression Network**

Module	GO Term	DAVID Score
1	Cell Morphology	2.18
2	Protein Transport	4.01
3	Intracellular Organelle Lumen	10.88
4	Innate Immune Response	10.51
5	Signaling	17.81
6	RNA binding	2.75
7	Translation	2.64
8	Zinc Binding	5.00
9	Tissue Morphogenesis	4.17
10	Oxidative Phosphorylation	2.22
11	Sarcomere	1.99
12	Mitochondrion	5.48
13	Transcription	1.95
14	Transit Peptide	8.04
15	Cytochrome P450	9.51
16	Immune Response	5.4
17	Lung Development	3.25
18	Zinc Binding	2.14
19	Contractile Fiber	2.09

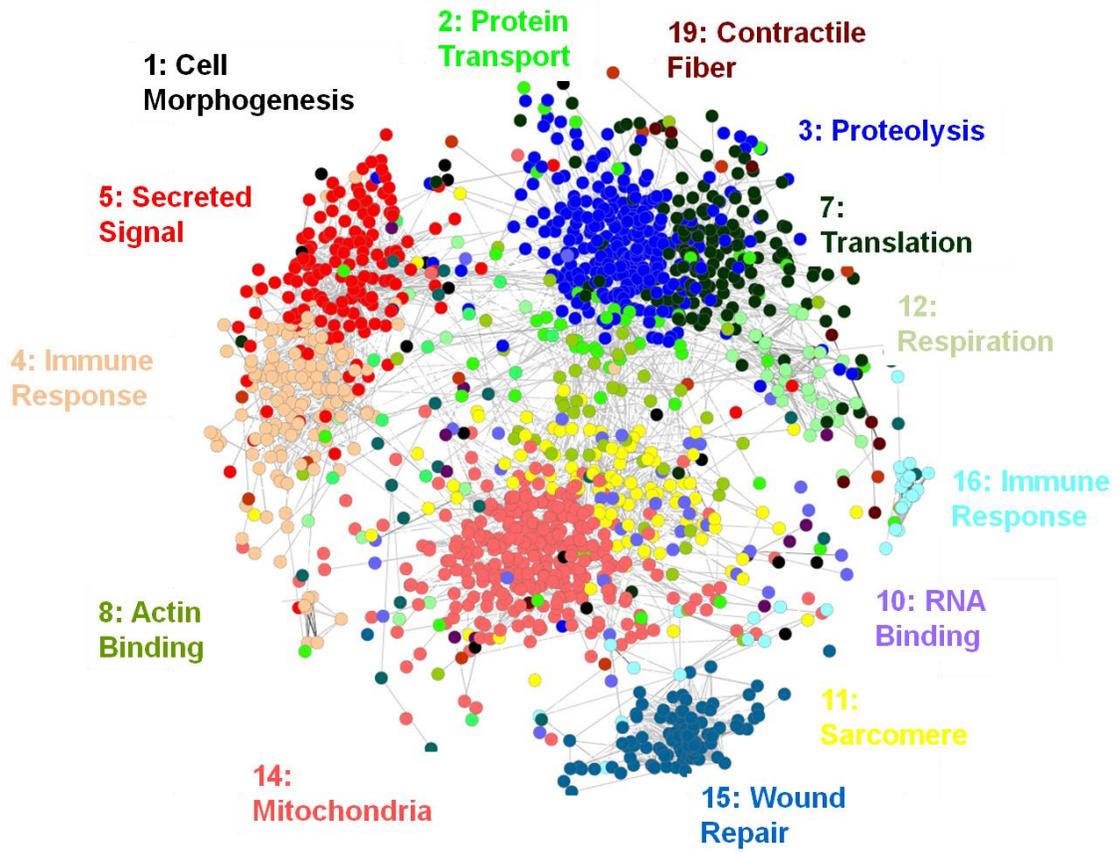


Figure 1.

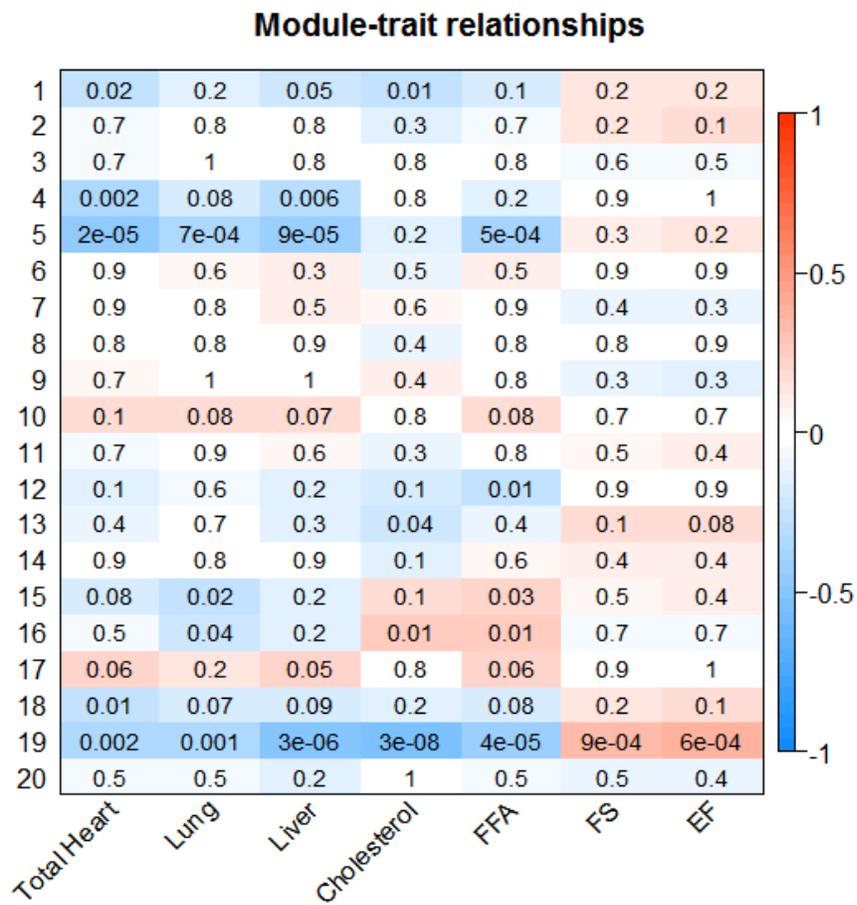
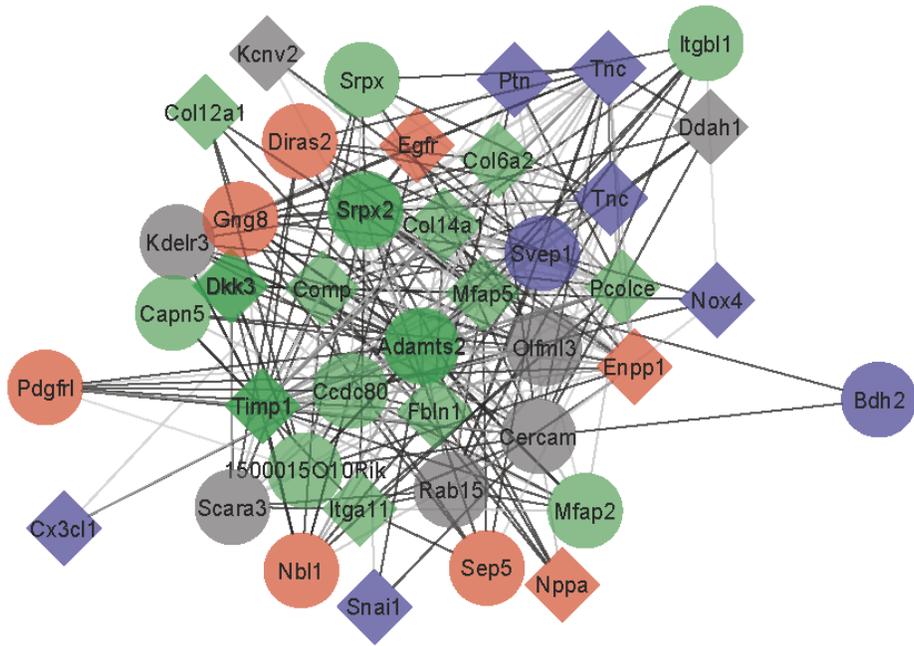


Figure 2.



**Figure 3.**



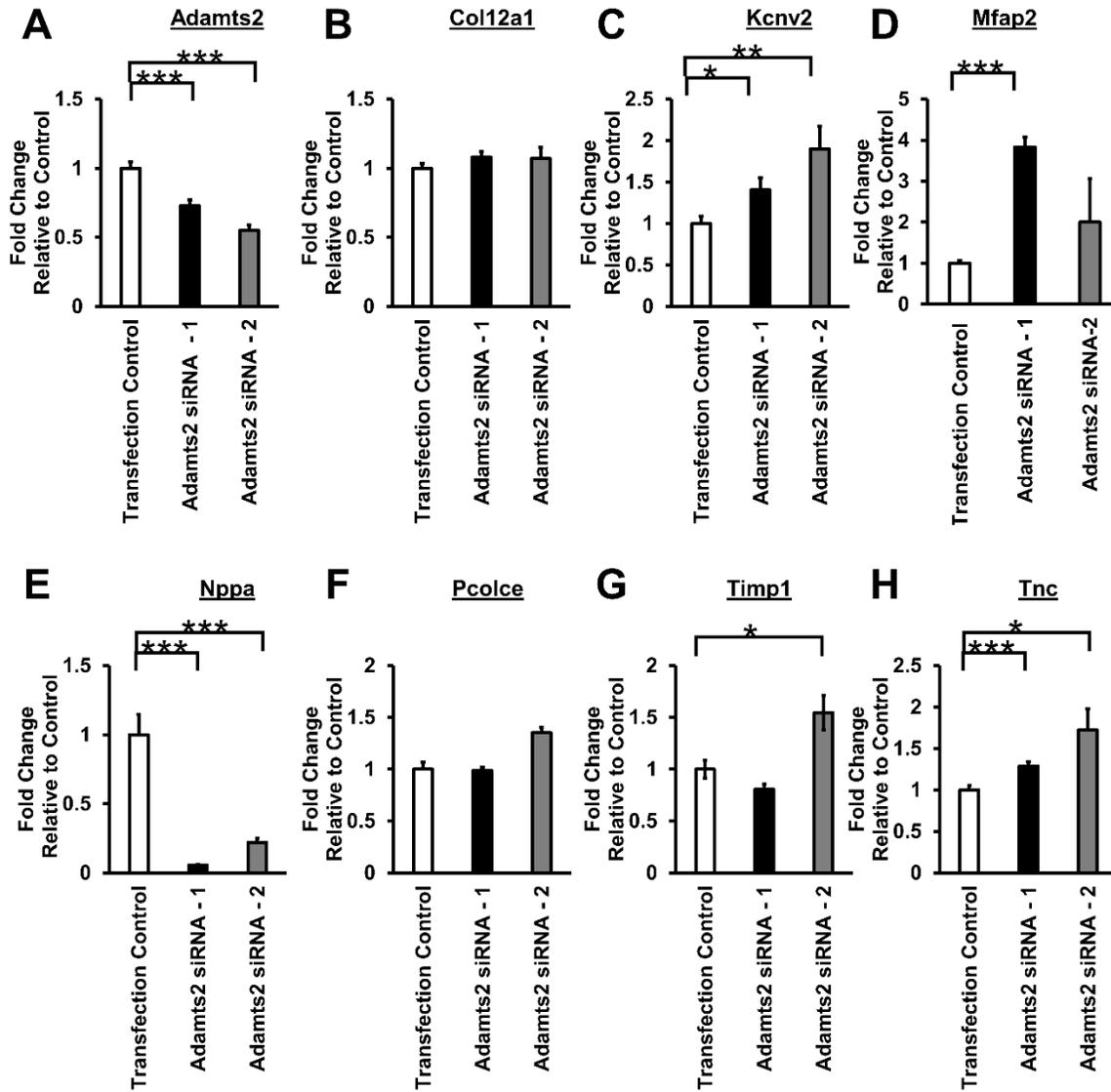


Figure 5.

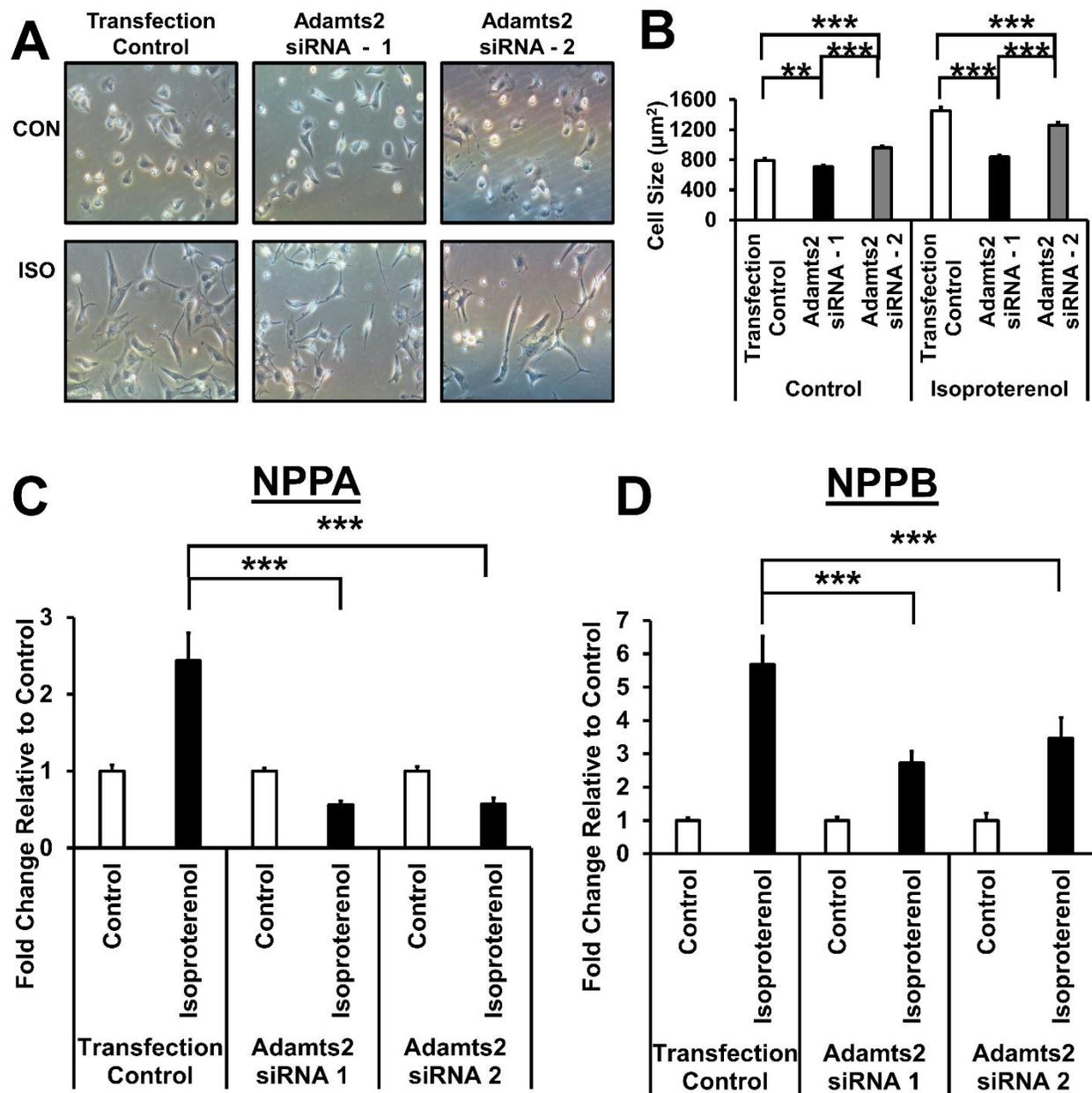


Figure 6.

**Supplemental Materials**

**Supplementary Table 1.** Strains used in this study

<b>Control Mice</b>	<b>Treated Mice</b>
129X1/SvJ	129X1/SvJ
A/J	A/J
AKR/J	AKR/J
AXB10/PgnJ	AXB10/PgnJ
AXB-12/PgnJ	
AXB18	AXB18
AXB19/PgnJ	AXB19/PgnJ
AXB-20/PgnJ	AXB-20/PgnJ
AXB-4/PgnJ	AXB-4/PgnJ
AXB-6/PgnJ	AXB-6/PgnJ
AXB8/PgnJ	AXB8/PgnJ
BALB/cByJ	BALB/cByJ
BALB/cJ	BALB/cJ
BTBRT<+>tf/J	BTBRT<+>tf/J
BUB/BnJ	BUB/BnJ
BXA-1/PgnJ	BXA-1/PgnJ
BXA-11/PgnJ	BXA-11/PgnJ
BXA-12/PgnJ	
BXA-14/PgnJ	BXA14/PgnJ
BXA16/PgnJ	BXA16/PgnJ
BXA-2/PgnJ	BXA-2/PgnJ
BXA24/PgnJ	BXA24/PgnJ
BXA-4/PgnJ	BXA-4/PgnJ
BXA-7/PgnJ	BXA-7/PgnJ
BXA-8/PgnJ	BXA-8/PgnJ
BXD-1/TyJ	
BXD-11/TyJ	BXD-11/TyJ
BXD-12/TyJ	BXD-12/TyJ
BXD-14/TyJ	BXD-14/TyJ
	BXD-19/TyJ
BXD21/TyJ	BXD21/TyJ
BXD-24/TyJ	BXD-24/TyJ
	BXD-31/TyJ
BXD32/TyJ	BXD32/TyJ
BXD-34/TyJ	
BXD-38/TyJ	BXD-38/TyJ

BXD39/TyJ	BXD39/TyJ
BXD40/TyJ	BXD40/TyJ
BXD43	BXD43
BXD44	BXD44
BXD45	BXD45
BXD48	BXD48
BXD49	BXD49
BXD-5/TyJ	BXD-5/TyJ
BXD50	BXD50
BXD55	BXD55
BXD56	BXD56
	BXD-6/TyJ
BXD61	BXD61
BXD62	BXD62
BXD64	BXD64
BXD66	BXD66
BXD68	BXD68
BXD69	
BXD70	BXD70
BXD71	BXD71
BXD73	BXD73
BXD74	BXD74
BXD75	BXD75
BXD79	BXD79
BXD84	BXD84
BXD85	BXD85
BXD86	
BXD87	BXD87
BXH-19/TyJ	BXH-19/TyJ
BXH-6/TyJ	BXH-6/TyJ
BXH-9/TyJ	BXH-9/TyJ
	BXHA1
BXHB2	BXHB2
C3H/HeJ	C3H/HeJ
C57BL/6J	C57BL/6J
C57BLKS/J	C57BLKS/J
CBA/J	C58/J
	CBA/J
CE/J	CE/J

CXB-11/HiAJ	CXB-11/HiAJ
CXB-12/HiAJ	CXB-12/HiAJ
CXB-13/HiAJ	CXB-13/HiAJ
CXB-3/ByJ	CXB-3/ByJ
CXB-6/ByJ	CXB-6/ByJ
CXB-7/ByJ	CXB-7/ByJ
CXBH	CXBH
DBA/2J	DBA/2J
FVB/NJ	FVB/NJ
KK/HiJ	KK/HiJ
LG/J	LG/J
LP/J	LP/J
	MA/MyJ
NOD/LtJ	NOD/LtJ
NON/LtJ	NON/LtJ
NZB/BINJ	NZB/BINJ
NZW/LacJ	NZW/LacJ
PL/J	PL/J
RIIS/J	RIIS/J
SEA/GnJ	SEA/GnJ
SJL/J	SJL/J
SM/J	SM/J
SWR/J	SWR/J

**Supplementary Table 2. Top Gene Ontologies for Modules of the Control Left Ventricular MICA Network**

Module	GO Term	DAVID Score
1	Protein Transport	2.76
2	Regulation of Transcription	2.44
3	Membrane	2.7
4	Regulation of Cytoskeleton Organization	1.94
5	mitochondrion	4.1
6	Extracellular Matrix	6.35
7	Chemotaxis	2.52
8	Secreted Signals	8.4
9	Mitochondrion	2.4
10	RNA-binding	5.19
11	Drug Response	8.11
12	Antigen Processing and Presentation	5.6
13	mitochondrion	7.29
14	Regulation of Apoptosis	1.05
15	PDZ Domain	1.7
16	Circadian Rhythm	2.82
17	angiogenesis	2.58
18	Regulation of Transcription	2.45
19	Protein Transport	5.45
20	Adaptive Immune Response	1.87

**Supplementary Table 3. Genes in Module 1 of the treated network with significant correlation to RV weight**

Gene Name	Correlation to RV Weight	Pvalue
<i>Mobk1b</i>	0.493743	5.7E-7
<i>Pfkp</i>	0.490219	7.1E-7
<i>Tcf21</i>	0.384174	1.6E-4
<i>Osbp2</i>	-0.38409	1.6E-4
<i>Pla1a</i>	0.373825	2.4E-4

**Supplementary Table 4. Top Gene Ontologies for Modules of the Ratio of Treated to Control Left Ventricular MICA Network**

Module	GO Term	DAVID Score
1	Mitochondrion	2.34
2	Protein catabolism	1.9
3	Mitochondrial inner membrane	3.71
4	Mitochondrion	12.14
5	Protein Transport	1.21
6	Mitochondrion	15.7
7	DNA repair	3
8	Arrhythmogenic RV cardiomyopathy	2.38
9	Protein Transport	3.23
10	transcription	3.81
11	membrane-enclosed lumen	12.03
12	contractile fiber	2.7
13	DNA Binding	4.47
14	Secreted Signal	18.57
15	RNA-binding	2.33
16	Ribosome	2.51
17	Endoplasmic reticulum	2.67
18	Protein Transport	4
19	mitochondrion	3.06
20	Ribosome	1.63

**Supplementary Table 5. Drivers and reactive genes in module 5**

Probe ID	Gene Symbol	Forward	Reverse	% Forward	Category
ILMN_1216661	COL6A2	19	0	100	Driver
ILMN_2698499	NOX4	1	0	100	Driver
ILMN_1226259	ADAMTS2	23	2	92	Driver
ILMN_2729103	ADAMTS2	24	3	89	Driver
ILMN_1238000	SRPX	7	1	88	Driver
ILMN_2671755	CERCAM	22	4	84	Driver
ILMN_1253741	PCOLCE	16	3	84	Driver
ILMN_1256676	DDAH1	5	1	83	Driver
ILMN_2728985	FBLN1	13	4	76	Driver
ILMN_1217009	RAB15	9	3	75	Driver
ILMN_1249000	1500015O10RIK	6	2	75	Driver
ILMN_1253791	KCNV2	2	1	67	Intermediary
ILMN_2981542	MFAP2	7	4	64	Intermediary
ILMN_2701778	ITGA11	3	2	60	Intermediary
ILMN_2591027	COL14A1	8	6	57	Intermediary
ILMN_2903926	PDGFRL	4	3	57	Intermediary
ILMN_2636424	ITGBL1	1	1	50	Intermediary
ILMN_2862538	COL12A1	2	2	50	Intermediary
ILMN_2926480	DIRAS2	5	5	50	Intermediary
ILMN_3128725	EGFR	2	2	50	Intermediary
ILMN_2463181	TNC	7	8	47	Intermediary
ILMN_2463180	TNC	6	8	43	Intermediary
ILMN_3161547	NPPA	2	3	40	Intermediary
ILMN_1234824	CCDC80	5	8	38	Intermediary
ILMN_1231851	ENPP1	5	9	36	Intermediary
ILMN_2627041	CX3CL1	1	2	33	Intermediary
ILMN_2627566	CAPN5	2	4	33	Intermediary
ILMN_2638114	PTN	2	4	33	Intermediary
ILMN_1259653	GNG8	3	6	33	Intermediary
ILMN_2675697	KDEL3	3	6	33	Intermediary
ILMN_2698728	SRPX2	5	10	33	Intermediary
ILMN_2818294	SRPX2	4	12	25	Reactive
ILMN_2886618	NBL1	1	5	17	Reactive
ILMN_1233455	OLFML3	1	13	7	Reactive
ILMN_2852957	DKK3	0	1	0	Reactive
ILMN_2746556	DKK3	0	2	0	Reactive
ILMN_2904765	SNAI1	0	3	0	Reactive
ILMN_2706268	SCARA3	0	6	0	Reactive
ILMN_2869110	COMP	0	9	0	Reactive
ILMN_1225835	MFAP5	0	13	0	Reactive
ILMN_3103896	TIMP1	0	15	0	Reactive
ILMN_2691951	SVEP1	0	16	0	Reactive

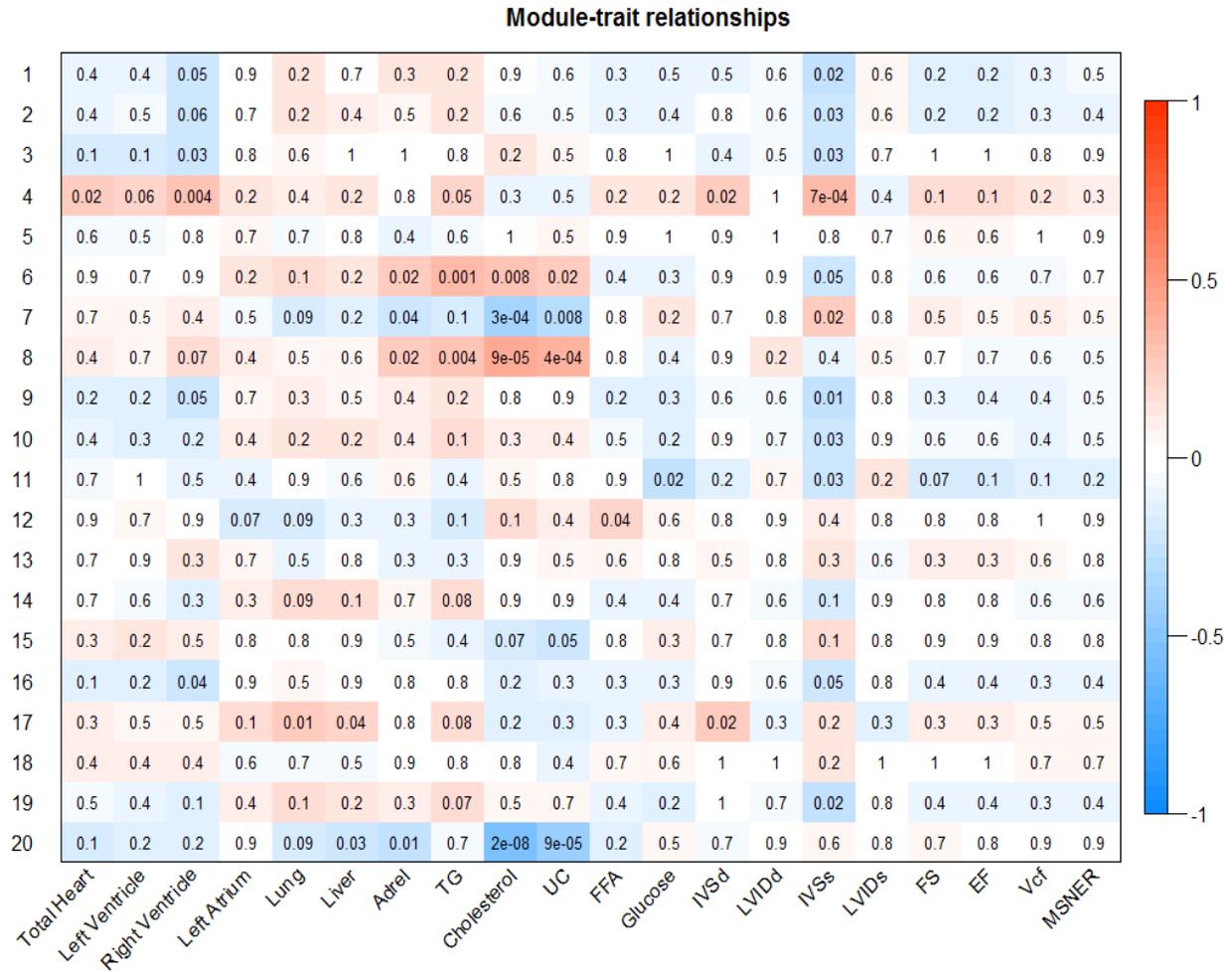
---

ILMN_2769918	TIMP1	0	18	0	Reactive
--------------	-------	---	----	---	----------

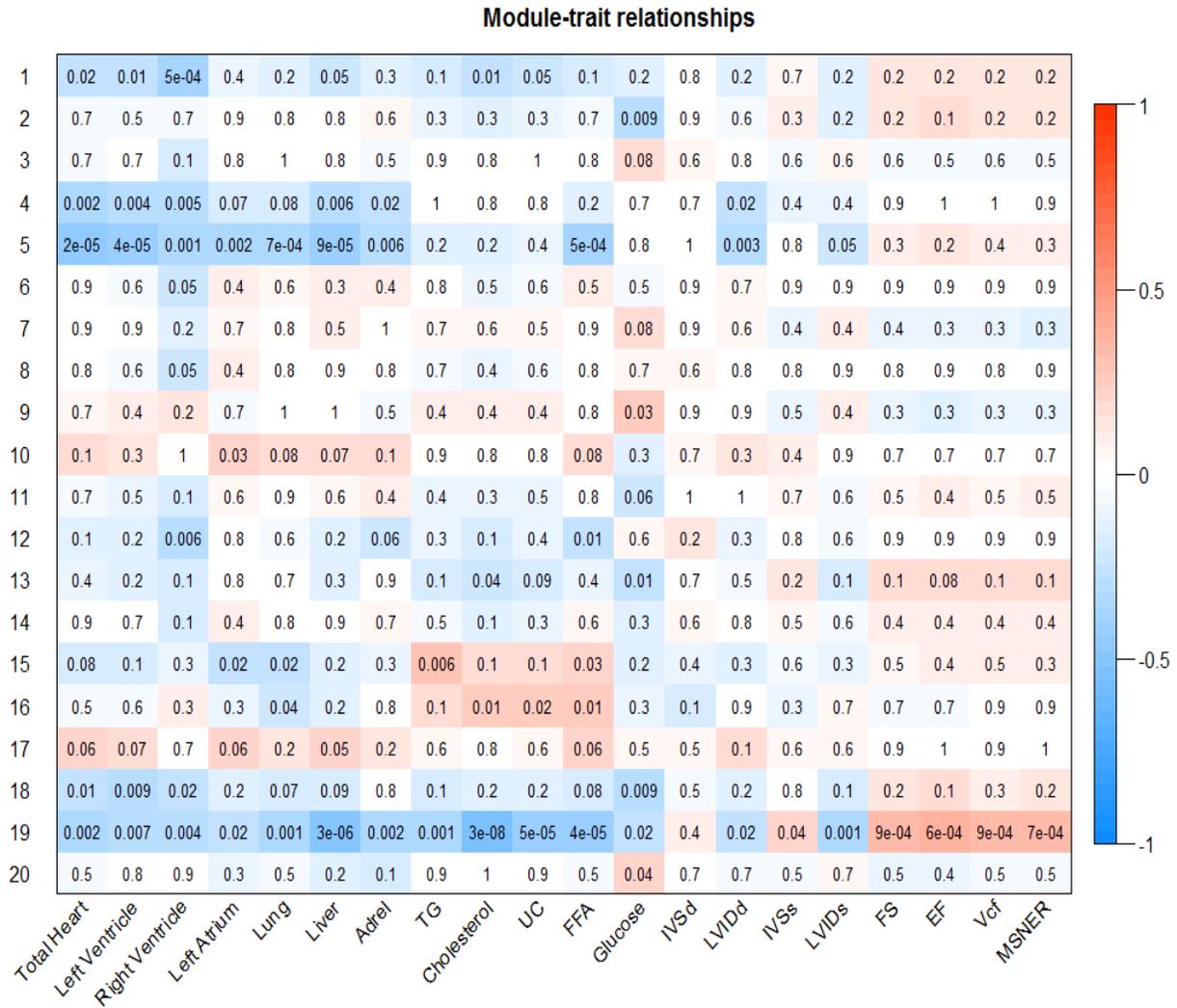
---

**Supplementary Table 6. Characteristics of Module 5 *Adamts2* target genes that were selected for *in vitro* validation using siRNA knockdown in NRVMs.**

	Gene Symbol	Probe ID	Predicted Strength of relationship to <i>Adamts2</i> by NEO	Number of Predicted Edges in Module 5	Previous Association with cardiovascular disease
1	<i>Col12a1</i>	ILMN_2862538	0.52919	4	Recessive and dominant mutations in COL12A1 cause a novel EDS/myopathy overlap syndrome in humans and mice
2	<i>Kcnv2</i>	ILMN_1253791	0.55139	3	
3	<i>Mfap2</i>	ILMN_2981542	0.54725	13	
4	<i>Nppa</i>	ILMN_3161547	0.5395	5	Hypertrophic marker
5	<i>Pcolce</i>	ILMN_1253741	0.75863 0.77621	19	Candidate gene for RV weight
6	<i>Timp1</i>	ILMN_3103896	0.72486 0.69964	15	ECM regulator that is increased during heart failure
		ILMN_2769918	0.73708 0.71147	18	
7	<i>Tnc</i>	ILMN_2463181	N/A	15	Expression of TNC has been suggested to accelerate adverse ventricular remodeling, cardiac failure, and fibrosis in the residual myocardium after MI.
		ILMN_2463180	0.62397 0.6225	14	

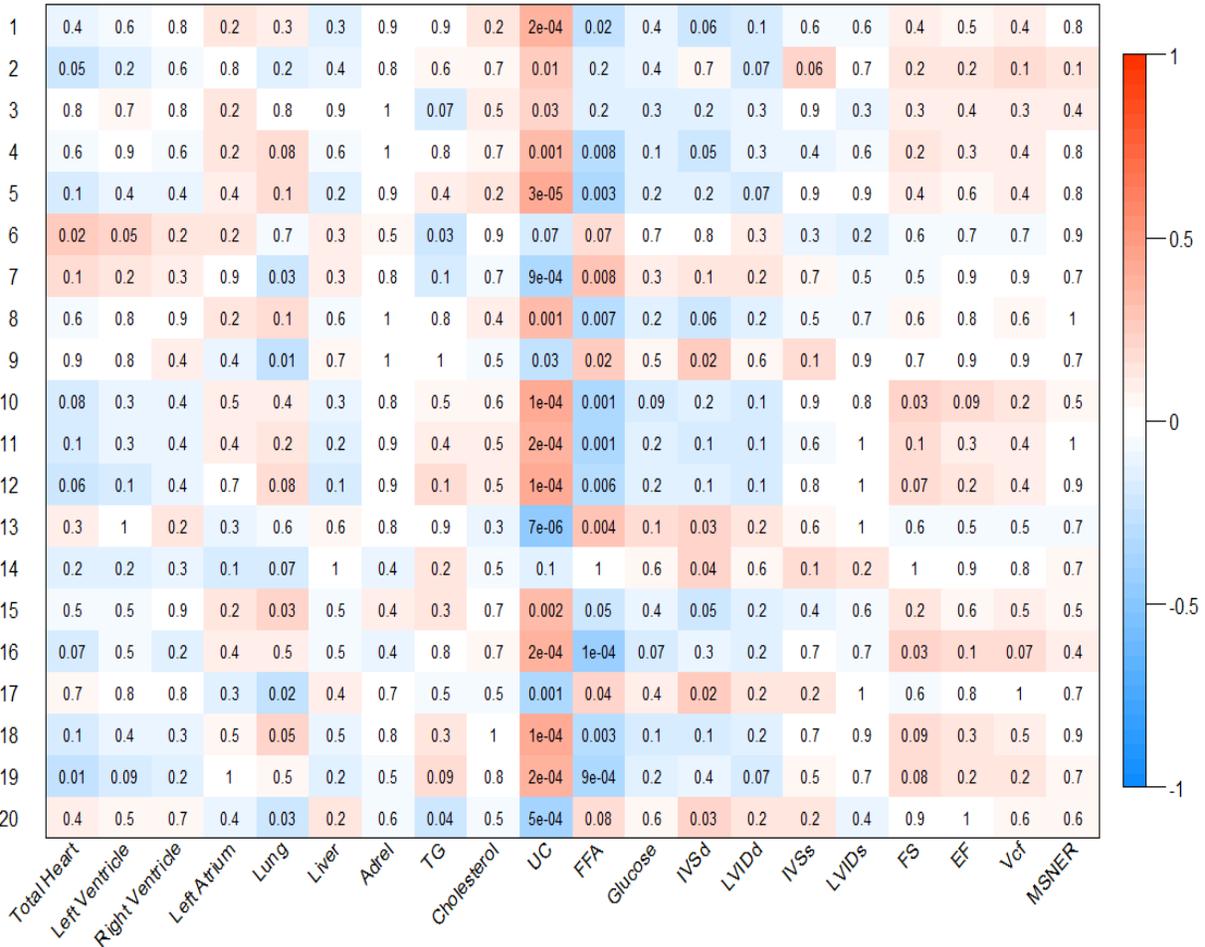


**Supplementary Figure 1. Module-Trait Correlation for Control Network.** A weighted PCA algorithm was used to calculate the first weighted principle component of each modules. These weighted first principle components were then correlated to the HF-related phenotypes. Strength of correlation is indicated by color, while the p-value of the correlation is indicated by the numbers in the table.

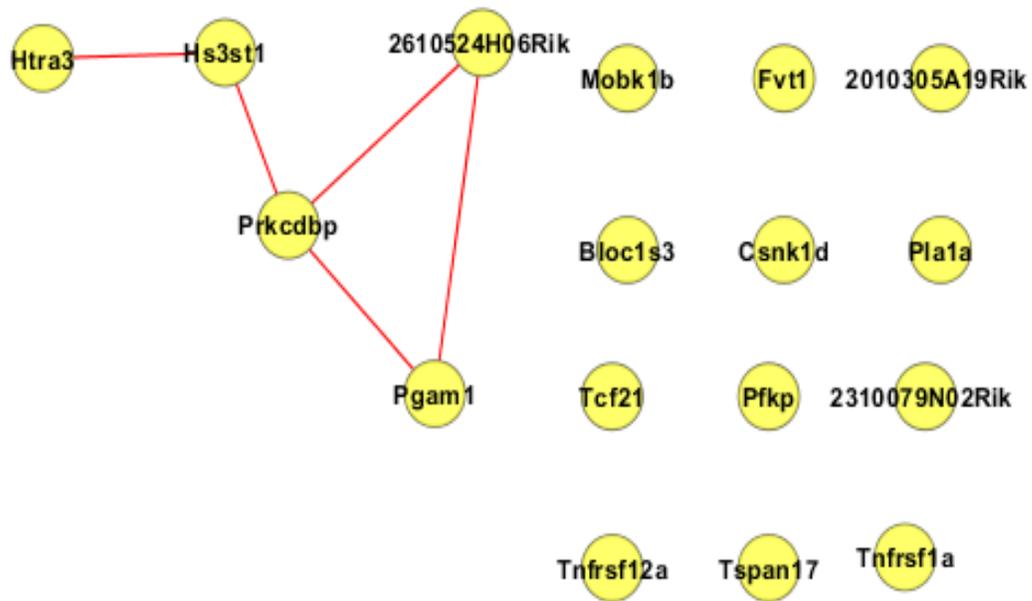


**Supplementary Figure 2. Module-Trait Correlation for Treated Network.** A weighted PCA algorithm was used to calculate the first weighted principle component of each modules. These weighted first principle components were then correlated to the HF-related phenotypes. Strength of correlation is indicated by color, while the p-value of the correlation is indicated by the numbers in the table.

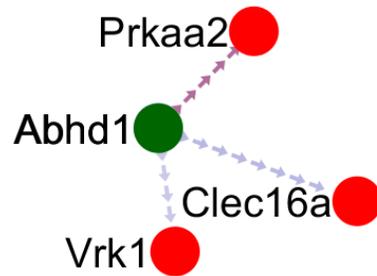
**Module-trait relationships**



**Supplementary Figure 3. Module-Trait Correlation for Induction Network.** A weighted PCA algorithm was used to calculate the first weighted principle component of each modules. These weighted first principle components were then correlated to the HF-related phenotypes. Strength of correlation is indicated by color, while the p-value of the correlation is indicated by the numbers in the table.

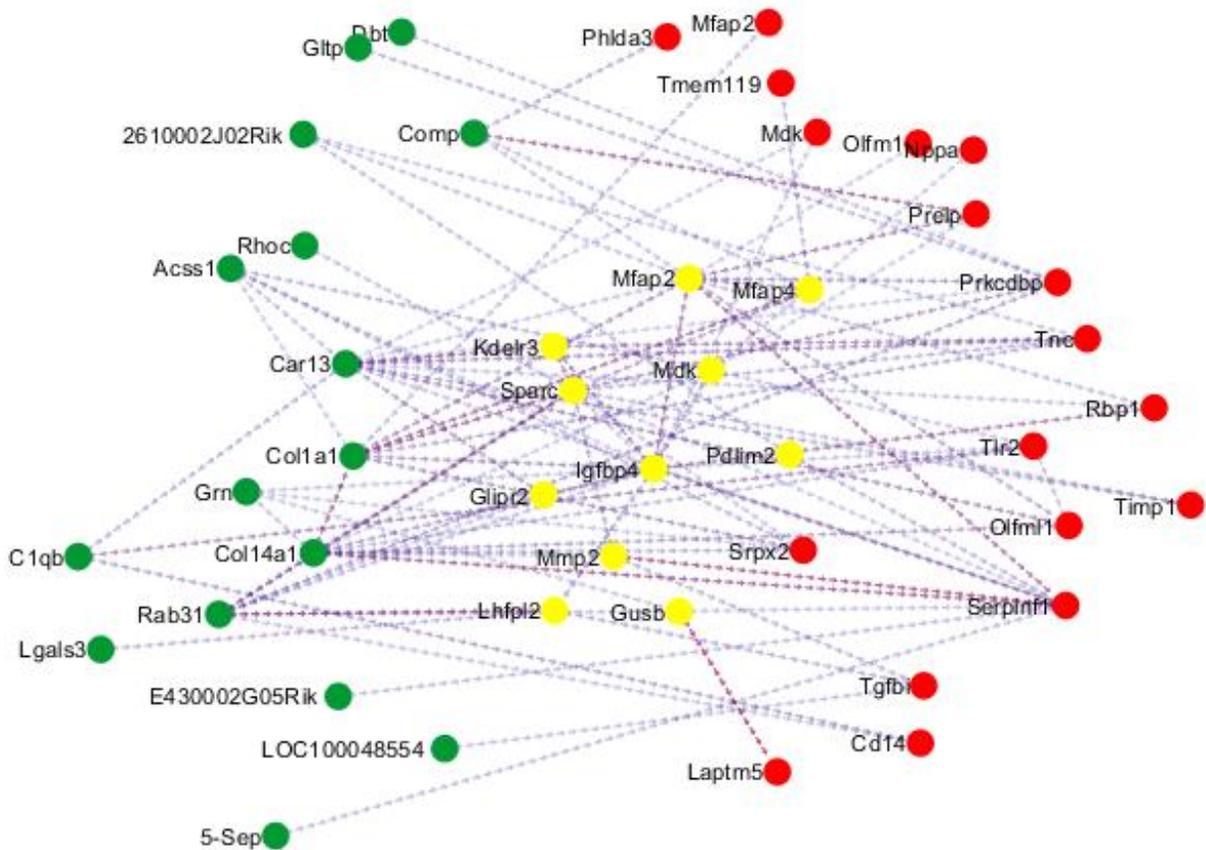


**Supplementary Figure 4. Graphical depiction of module 1 in the control network.** The probes depicted here are the 'core' probes with module memberships within the module of greater than 50%. Only 5 of these genes are connected to one another with significant MIC edges, suggesting that the module as a whole has poor underlying structure and is not suitable for further analyses, such as NEO.



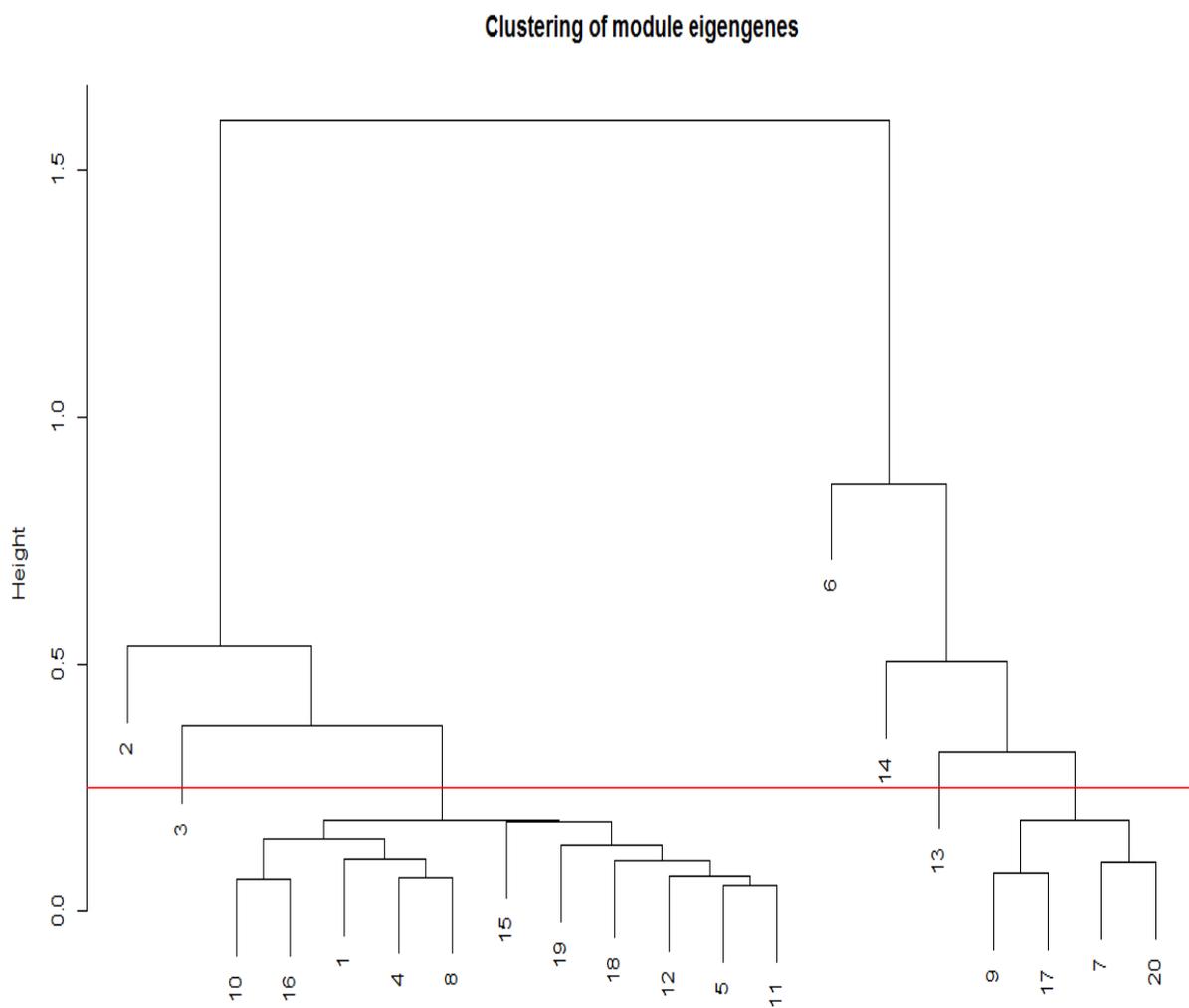
**Supplementary Figure 5. Near Edge Orientation of Module 19 of the treated network.**

Arrows indicate the direction of the interaction of the probes in the module, while the color of the edge represents the strength of that interaction, with blue indicating a lower MIC score and red indicating a high MIC score. Nodes are colored based on their status: Green indicates drivers, where 75% or more of the edges for the probe travel away from the node in question. Red indicates reactive genes, where 75% or more of the edges for that probe travel towards the node. In this case, it is clear that Abhd1 is the driver for all of the directed edges in the module and affects the other three linked genes in turn.



**Supplementary Figure 6. Near Edge Orientation of Module 14 of the delta network.**

Arrows indicate the direction of the interaction of the probes in the module, while the color of the edge represents the strength of that interaction, with blue indicating a lower MIC score and red indicating a high MIC score. Nodes are colored based on their status: Green indicates drivers, where 75% or more of the edges for the probe travel away from the node in question. Red indicates reactive genes, where 75% or more of the edges for that probe travel towards the node and yellow are genes which meet neither of these criteria. Comparison of this module to the NEO performed on module 5 of the treated network reveals that module 14 does not possess a clear central driver (similar to *Pcolce* or *Adamts2* in module 5) nor reactive hub (similar to *Timp1* in module 5).



**Supplementary Figure 7. Hierarchical Clustering of Induction Module Eigengenes.**

Heirarchical clustering of the eigengenes from the Induction dataset reveals that many of these modules are highly correlated to one another with  $R > 0.8$  (red line)

## References

- Albanese, D., Filosi, M., Visintainer, R., Riccadonna, S., Jurman, G., and Furlanello, C. (2013). Minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers. *Bioinformatics* 29, 407–408.
- Arai, K., Kasashima, Y., Kobayashi, A., Kuwano, A., and Yoshihara, T. (2002). TGF-beta alters collagen XII and XIV mRNA levels in cultured equine tenocytes. *Matrix Biol.* 21, 243–250.
- Aten, J.E., Fuller, T.F., Lusic, A.J., and Horvath, S. (2008). Using genetic markers to orient the edges in quantitative trait networks: the NEO software. *BMC Syst. Biol.* 2, 34.
- Bao, M.-W., Cai, Z., Zhang, X.-J., Li, L., Liu, X., Wan, N., Hu, G., Wan, F., Zhang, R., Zhu, X., et al. (2015). Dickkopf-3 protects against cardiac dysfunction and ventricular remodelling following myocardial infarction. *Basic Res. Cardiol.* 110, 25.
- Barbier, M., Gross, M.-S., Aubart, M., Hanna, N., Kessler, K., Guo, D.-C., Tosolini, L., Ho-Tin-Noe, B., Regalado, E., Varret, M., et al. (2014). MFAP5 loss-of-function mutations underscore the involvement of matrix alteration in the pathogenesis of familial thoracic aortic aneurysms and dissections. *Am. J. Hum. Genet.* 95, 736–743.
- Barp, A., Bello, L., Politano, L., Melacini, P., Calore, C., Polo, A., Vianello, S., Sorarù, G., Semplicini, C., Pantic, B., et al. (2015). Genetic Modifiers of Duchenne Muscular Dystrophy and Dilated Cardiomyopathy. *PLoS One* 10, e0141240.
- Barton, P.J.R., Birks, E.J., Felkin, L.E., Cullen, M.E., Koban, M.U., and Yacoub, M.H. (2003). Increased expression of extracellular matrix regulators TIMP1 and MMP1 in deteriorating heart failure. *J. Hear. Lung Transplant. Off. Publ. Int. Soc. Hear. Transplant.* 22, 738–744.
- Bekhouche, M., Leduc, C., Dupont, L., Janssen, L., Delolme, F., Vadon-Le Goff, S., Smargiasso, N., Baiwir, D., Mazzucchelli, G., Zanella-Cleon, I., et al. (2016). Determination of the substrate

repertoire of ADAMTS2, 3, and 14 significantly broadens their functions and identifies extracellular matrix organization and TGF- $\beta$  signaling as primary targets. *FASEB J.*

Bennett, B.J., Farber, C.R., Orozco, L., Kang, H.M., Ghazalpour, A., Siemers, N., Neubauer, M., Neuhaus, I., Yordanova, R., Guan, B., et al. (2010). A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res.* *20*, 281–290.

Bennett, B.J., Davis, R.C., Civelek, M., Orozco, L., Wu, J., Qi, H., Pan, C., Packard, R.R.S., Eskin, E., Yan, M., et al. (2015). Genetic Architecture of Atherosclerosis in Mice: A Systems Genetics Analysis of Common Inbred Strains. *PLoS Genet.* *11*, e1005711.

Brown, D. a, Beygui, R.E., MacLellan, W.R., Laks, H., Dunn, J.C.Y., and Wu, B.M. (2005). Modulation of gene expression in neonatal rat cardiomyocytes by surface modification of polylactide-co-glycolide substrates. *J. Biomed. Mater. Res. A* *74*, 419–429.

Calabrese, G., Bennett, B.J., Orozco, L., Kang, H.M., Eskin, E., Dombret, C., De Backer, O., Lusic, A.J., and Farber, C.R. (2012). Systems genetic analysis of osteoblast-lineage cells. *PLoS Genet.* *8*, e1003150.

Camargo, A., and Azuaje, F. (2007). Linking gene expression and functional network data in human heart failure. *PLoS One* *2*, e1347.

Chen, L., Ge, Q., Black, J.L., Deng, L., Burgess, J.K., and Oliver, B.G.G. (2013). Differential regulation of extracellular matrix and soluble fibulin-1 levels by TGF- $\beta_1$  in airway smooth muscle cells. *PLoS One* *8*, e65544.

Chessel, D., Dufour, A.B., and Thioulouse, J. (2004). The ade4 package - I : One-table methods. *4*, 5–10.

Cooley, M.A., Fresco, V.M., Dorlon, M.E., Twal, W.O., Lee, N. V, Barth, J.L., Kern, C.B., Iruela-Arispe, M.L., and Argraves, W.S. (2012). Fibulin-1 is required during cardiac ventricular

morphogenesis for versican cleavage, suppression of ErbB2 and Erk1/2 activation, and to attenuate trabecular cardiomyocyte proliferation. *Dev. Dyn.* 241, 303–314.

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R. a (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4, P3.

Dewey, F.E., Perez, M. V, Wheeler, M.T., Watt, C., Spin, J., Langfelder, P., Horvath, S., Hannenhalli, S., Cappola, T.P., and Ashley, E. a (2011). Gene coexpression network topology of cardiac development, hypertrophy, and failure. *Circ. Cardiovasc. Genet.* 4, 26–35.

Dobaczewski, M., Chen, W., and Frangogiannis, N.G. (2011). Transforming growth factor (TGF)- $\beta$  signaling in cardiac remodeling. *J. Mol. Cell. Cardiol.* 51, 600–606.

Drozдов, I., Didangelos, A., Yin, X., Zampetaki, A., Abonnenc, M., Murdoch, C., Zhang, M., Ouzounis, C. a, Mayr, M., Tsoka, S., et al. (2013). Gene network and proteomic analyses of cardiac responses to pathological and physiological stress. *Circ. Cardiovasc. Genet.* 6, 588–597.

Ehnis, T., Dieterich, W., Bauer, M., Kresse, H., and Schuppan, D. (1997). Localization of a binding site for the proteoglycan decorin on collagen XIV (undulin). *J. Biol. Chem.* 272, 20414–20419.

Farber, C.R. (2013). Systems-level analysis of genome-wide association data. *G3 (Bethesda)*. 3, 119–129.

Farber, C.R., Aten, J.E., Farber, E. a, de Vera, V., Gularte, R., Islas-Trejo, A., Wen, P., Horvath, S., Lucero, M., Lusis, A.J., et al. (2009). Genetic dissection of a major mouse obesity QTL (Carfhg2): integration of gene expression and causality modeling. *Physiol. Genomics* 37, 294–302.

Farber, C.R., Bennett, B.J., Orozco, L., Zou, W., Lira, A., Kostem, E., Kang, H.M., Furlotte, N., Berberyan, A., Ghazalpour, A., et al. (2011). Mouse genome-wide association and systems genetics identify *Asx12* as a regulator of bone mineral density and osteoclastogenesis. *PLoS Genet.* 7, e1002038.

Frangogiannis, N.G. (2012). Regulation of the inflammatory response in cardiac repair. *Circ. Res.* 110, 159–173.

Ghazalpour, A., Rau, C.D., Farber, C.R., Bennett, B.J., Orozco, L.D., van Nas, A., Pan, C., Allayee, H., Beaven, S.W., Civelek, M., et al. (2012). Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm. Genome* 23, 680–692.

Goding, J.W., Grobden, B., and Slegers, H. (2003). Physiological and pathophysiological functions of the ecto-nucleotide pyrophosphatase/phosphodiesterase family. *Biochim. Biophys. Acta* 1638, 1–19.

Grossman, T.R., Gamliel, A., Wessells, R.J., Taghli-Lamalle, O., Jepsen, K., Ocorr, K., Korenberg, J.R., Peterson, K.L., Rosenfeld, M.G., Bodmer, R., et al. (2011). Over-expression of *DSCAM* and *COL6A2* cooperatively generates congenital heart defects. *PLoS Genet.* 7, e1002344.

Gu, X., Xu, J., Yang, X.-P., Peterson, E., and Harding, P. (2015). Fractalkine neutralization improves cardiac function after myocardial infarction. *Exp. Physiol.* 100, 805–817.

Hall, M.-C., Young, D.A., Waters, J.G., Rowan, A.D., Chantry, A., Edwards, D.R., and Clark, I.M. (2003). The comparative role of activator protein 1 and Smad factors in the regulation of *Timp-1* and *MMP-1* gene expression by transforming growth factor-beta 1. *J. Biol. Chem.* 278, 10304–10313.

Hanssen, E., Hew, F.H., Moore, E., and Gibson, M.A. (2004). MAGP-2 has multiple binding regions on fibrillins and has covalent periodic association with fibrillin-containing microfibrils. *J. Biol. Chem.* 279, 29185–29194.

Hessel, M., Steendijk, P., den Adel, B., Schutte, C., and van der Laarse, A. (2009). Pressure overload-induced right ventricular failure is associated with re-expression of myocardial tenascin-C and elevated plasma tenascin-C levels. *Cell. Physiol. Biochem.* 24, 201–210.

Hofmann Bowman, M.A., and McNally, E.M. (2012). Genetic pathways of vascular calcification. *Trends Cardiovasc. Med.* 22, 93–98.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.

Hui, S.T., Parks, B.W., Org, E., Norheim, F., Che, N., Pan, C., Castellani, L.W., Charugundla, S., Dirks, D.L., Psychogios, N., et al. (2015). The genetic architecture of NAFLD among inbred strains of mice. *Elife* 4, e05607.

Hung, W.-T., Wu, F.-J., Wang, C.-J., and Luo, C.-W. (2012). DAN (NBL1) specifically antagonizes BMP2 and BMP4 and modulates the actions of GDF9, BMP2, and BMP4 in the rat ovary. *Biol. Reprod.* 86, 158, 1–9.

Ikonomidis, J.S., Hendrick, J.W., Parkhurst, A.M., Herron, A.R., Escobar, P.G., Dowdy, K.B., Stroud, R.E., Hapke, E., Zile, M.R., and Spinale, F.G. (2005). Accelerated LV remodeling after myocardial infarction in TIMP-1-deficient mice: effects of exogenous MMP inhibition. *Am. J. Physiol. Hear. Circ. Physiol.* 288, H149–H158.

Jinnin, M., Ihn, H., Asano, Y., Yamane, K., Trojanowska, M., and Tamaki, K. (2004). Tenascin-C upregulation by transforming growth factor-beta in human dermal fibroblasts involves Smad3, Sp1, and Ets1. *Oncogene* 23, 1656–1667.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118–127.

Kessler-Icekson, G., Schlesinger, H., Freimann, S., and Kessler, E. (2006). Expression of procollagen C-proteinase enhancer-1 in the remodeling rat heart is stimulated by aldosterone. *Int. J. Biochem. Cell Biol.* 38, 358–365.

Kosla, J., Dvorak, M., and Cermak, V. (2013). Molecular analysis of the TGF-beta controlled gene expression program in chicken embryo dermal myofibroblasts. *Gene* 513, 90–100.

Kuroda, J., Ago, T., Matsushima, S., Zhai, P., Schneider, M.D., and Sadoshima, J. (2010). NADPH oxidase 4 (Nox4) is a major source of oxidative stress in the failing heart. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15565–15570.

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

Li, J., Wei, H., Chesley, A., Moon, C., Krawczyk, M., Volkova, M., Ziman, B., Margulies, K.B., Talan, M., Crow, M.T., et al. (2007). The pro-angiogenic cytokine pleiotrophin potentiates cardiomyocyte apoptosis through inhibition of endogenous AKT/PKB activity. *J. Biol. Chem.* 282, 34984–34993.

Li, X.-Q., Du, X., Li, D.-M., Kong, P.-Z., Sun, Y., Liu, P.-F., Wang, Q.-S., and Feng, Y.-M. (2015). ITGBL1 Is a Runx2 Transcriptional Target and Promotes Breast Cancer Bone Metastasis by Activating the TGF $\beta$  Signaling Pathway. *Cancer Res.* 75, 3302–3313.

McNally, E.M., Barefield, D.Y., and Puckelwartz, M.J. (2015). The genetic landscape of cardiomyopathy and its role in heart failure. *Cell Metab.* 21, 174–182.

Messaoudi, S., Zhang, A. Di, Griol-Charhbili, V., Escoubet, B., Sadoshima, J., Farman, N., and Jaisser, F. (2012). The epidermal growth factor receptor is involved in angiotensin II but not aldosterone/salt-induced cardiac remodelling. *PLoS One* 7, e30156.

Miljkovic-Licina, M., Hammel, P., Garrido-Urbani, S., Bradfield, P.F., Szepetowski, P., and Imhof, B.A. (2009). Sushi repeat protein X-linked 2, a novel mediator of angiogenesis. *FASEB J.* 23, 4105–4116.

Miljkovic-Licina, M., Hammel, P., Garrido-Urbani, S., Lee, B.P.-L., Meguenani, M., Chaabane, C., Bochaton-Piallat, M.-L., and Imhof, B.A. (2012). Targeting olfactomedin-like 3 inhibits tumor growth by impairing angiogenesis and pericyte coverage. *Mol. Cancer Ther.* 11, 2588–2599.

Milting, H., Ellinghaus, P., Seewald, M., Cakar, H., Bohms, B., Kassner, A., Körfer, R., Klein, M., Krahn, T., Kruska, L., et al. (2008). Plasma biomarkers of myocardial fibrosis and remodeling in terminal heart failure patients supported by mechanical circulatory support devices. *J. Heart Lung Transplant.* 27, 589–596.

Moali, C., Font, B., Ruggiero, F., Eichenberger, D., Rousselle, P., François, V., Oldberg, A., Bruckner-Tuderman, L., and Hulmes, D.J.S. (2005). Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1. *J. Biol. Chem.* 280, 24188–24194.

Moreno-Moral, A., Mancini, M., D'Amati, G., Camici, P., and Petretto, E. (2013). Transcriptional network analysis for the regulation of left ventricular hypertrophy and microvascular remodeling. *J. Cardiovasc. Transl. Res.* 6, 931–944.

Mudd, J.O., and Kass, D.A. (2008). Tackling heart failure in the twenty-first century. *Nature* 451, 919–928.

Nishioka, T., Onishi, K., Shimojo, N., Nagano, Y., Matsusaka, H., Ikeuchi, M., Ide, T., Tsutsui, H., Hiroe, M., Yoshida, T., et al. (2010). Tenascin-C may aggravate left ventricular remodeling and function after myocardial infarction in mice. *Am. J. Physiol. Heart Circ. Physiol.* 298, H1072–H1078.

Orozco, L.D., Bennett, B.J., Farber, C.R., Ghazalpour, A., Pan, C., Che, N., Wen, P., Qi, H.X., Mutukulu, A., Siemers, N., et al. (2012). Unraveling Inflammatory Responses using Systems Genetics and Gene-Environment Interactions in Macrophages. *Cell* 151, 658–670.

Park, C.C., Gale, G.D., de Jong, S., Ghazalpour, A., Bennett, B.J., Farber, C.R., Langfelder, P., Lin, A., Khan, A.H., Eskin, E., et al. (2011). Gene networks associated with conditional fear in mice identified using a systems genetics approach. *BMC Syst. Biol.* 5, 43.

Parkkinen, J. a, and Kaski, S. (2010). Searching for functional gene modules with interaction component models. *BMC Syst. Biol.* 4, 4.

Parks, B.W., Nam, E., Org, E., Kostem, E., Norheim, F., Hui, S.T., Pan, C., Civelek, M., Rau, C.D., Bennett, B.J., et al. (2013). Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab.* 17, 141–152.

Parks, B.W., Sallam, T., Mehrabian, M., Psychogios, N., Hui, S.T., Norheim, F., Castellani, L.W., Rau, C.D., Pan, C., Phun, J., et al. (2015). Genetic Architecture of Insulin Resistance in the Mouse. *Cell Metab.* 21, 334–346.

Porter, S., Clark, I.M., Kevorkian, L., and Edwards, D.R. (2005). The ADAMTS metalloproteinases. *Biochem. J.* 386, 15–27.

Rau, C.D., Wisniewski, N., Orozco, L.D., Bennett, B., Weiss, J., and Lusis, A.J. (2013). Maximal information component analysis: a novel non-linear network analysis method. *Front. Genet.* 4, 28.

Rau, C.D., Lusic, A.J., and Wang, Y. (2015a). Genetics of common forms of heart failure: challenges and potential solutions. *Curr. Opin. Cardiol.* 30, 222–227.

Rau, C.D., Wang, J., Avetisyan, R., Romay, M.C., Martin, L., Ren, S., Wang, Y., and Lusic, a. J. (2015b). Mapping Genetic Contributions to Cardiac Pathology Induced by Beta-Adrenergic Stimulation in Mice. *Circ. Cardiovasc. Genet.* 8, 40–49.

Reshef, D.N., Reshef, Y. a., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M., and Sabeti, P.C. (2011). Detecting Novel Associations in Large Data Sets. *Science* (80-. ). 334, 1518–1524.

Smyth, G.K. (2005). No Title. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, R. Gentleman, W. Irizarry, and W. Huber, eds. (New York: Springer), pp. 397–420.

Sun, M., Ouzounian, M., de Couto, G., Chen, M., Yan, R., Fukuoka, M., Li, G., Moon, M., Liu, Y., Gramolini, A., et al. (2013). Cathepsin-L ameliorates cardiac hypertrophy through activation of the autophagy-lysosomal dependent protein processing pathways. *J. Am. Heart Assoc.* 2, e000191.

Talior-Volodarsky, I., Connelly, K.A., Arora, P.D., Gullberg, D., and McCulloch, C.A. (2012).  $\alpha$ 11 integrin stimulates myofibroblast differentiation in diabetic cardiomyopathy. *Cardiovasc. Res.* 96, 265–275.

Tanaka, K., Arao, T., Maegawa, M., Matsumoto, K., Kaneda, H., Kudo, K., Fujita, Y., Yokote, H., Yanagihara, K., Yamada, Y., et al. (2009). SRPX2 is overexpressed in gastric cancer and promotes cellular migration and adhesion. *Int. J. Cancer* 124, 1072–1080.

Tao, G., Levay, A.K., Gridley, T., and Lincoln, J. (2011). Mmp15 is a direct target of Snai1 during endothelial to mesenchymal transformation and endocardial cushion development. *Dev. Biol.* 359, 209–221.

Tao, G., Levay, A.K., Peacock, J.D., Huk, D.J., Both, S.N., Purcell, N.H., Pinto, J.R., Galantowicz, M.L., Koch, M., Lucchesi, P.A., et al. (2012). Collagen XIV is important for growth and structural integrity of the myocardium. *J. Mol. Cell. Cardiol.* 53, 626–638.

Tillet, E., Wiedemann, H., Golbik, R., Pan, T.C., Zhang, R.Z., Mann, K., Chu, M.L., and Timpl, R. (1994). Recombinant expression and structural and binding properties of alpha 1(VI) and alpha 2(VI) chains of human collagen type VI. *Eur. J. Biochem.* 221, 177–185.

Tortorella, M.D., Malfait, F., Barve, R.A., Shieh, H.-S., and Malfait, A.-M. (2009). A review of the ADAMTS family, pharmaceutical targets of the future. *Curr. Pharm. Des.* 15, 2359–2374.

Uys, G.M., Ramburan, A., Loos, B., Kinnear, C.J., Korkie, L.J., Mouton, J., Riedemann, J., and Moolman-Smook, J.C. (2011). Myomegalin is a novel A-kinase anchoring protein involved in the phosphorylation of cardiac myosin binding protein C. *BMC Cell Biol.* 12, 18.

Vicente-Steijn, R., Scherptong, R.W.C., Kruithof, B.P.T., Duim, S.N., Goumans, M.J.T.H., Wisse, L.J., Zhou, B., Pu, W.T., Poelmann, R.E., Schaliij, M.J., et al. (2015). Regional differences in WT-1 and Tcf21 expression during ventricular development: implications for myocardial compaction. *PLoS One* 10, e0136025.

Weiss, J.N., Karma, A., MacLellan, W.R., Deng, M., Rau, C.D., Rees, C.M., Wang, J., Wisniewski, N., Eskin, E., Horvath, S., et al. (2012). “Good enough solutions” and the genetics of complex diseases. *Circ. Res.* 111, 493–504.

Wheeler, F.C., Tang, H., Marks, O. a, Hadnott, T.N., Chu, P.-L., Mao, L., Rockman, H. a, and Marchuk, D. a (2009). Tnni3k modifies disease progression in murine models of cardiomyopathy. *PLoS Genet.* 5, e1000647.

Wittköpper, K., Fabritz, L., Neef, S., Ort, K.R., Grefe, C., Unsöld, B., Kirchhof, P., Maier, L.S., Hasenfuss, G., Dobrev, D., et al. (2010). Constitutively active phosphatase inhibitor-1 improves cardiac contractility in young mice but is deleterious after catecholaminergic stress and with aging. *J. Clin. Invest.* 120, 617–626.

Xiong, J., Cui, X., Yuan, X., Yu, X., Sun, J., and Gong, Q. (2016). The Hippo/STE20 homolog SIK1 interacts with MOB1 to regulate cell proliferation and cell expansion in Arabidopsis. *J. Exp. Bot.* 67, 1461–1475.

Xuan, W., Liao, Y., Chen, B., Huang, Q., Xu, D., Liu, Y., Bin, J., and Kitakaze, M. (2011). Detrimental effect of fractalkine on myocardial ischaemia and heart failure. *Cardiovasc. Res.* 92, 385–393.

Yan, F., Wang, Y., Wu, X., Peshavariya, H.M., Dusting, G.J., Zhang, M., and Jiang, F. (2014). Nox4 and redox signaling mediate TGF- $\beta$ -induced endothelial cell apoptosis and phenotypic switch. *Cell Death Dis.* 5, e1010.

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4, Article17.

Zhang, Y., Liu, Y., Zhu, X.-H., Zhang, X.-D., Jiang, D.-S., Bian, Z.-Y., Zhang, X.-F., Chen, K., Wei, X., Gao, L., et al. (2014). Dickkopf-3 attenuates pressure overload-induced cardiac remodelling. *Cardiovasc. Res.* 102, 35–45.

## **Chapter 6**

### **Concluding Remarks**

In my dissertation work I present four distinct research projects involving the application of systems genetics approaches to understanding and characterizing molecular mechanisms in cardiovascular diseases. Cardiovascular diseases (CVDs) represent growing public health epidemic in the 21<sup>st</sup> century. Despite improvements in medical care, CVDs account for >17 million deaths globally each year, 80% of which occur in low-income and middle-income countries<sup>1</sup>.

### **Atherosclerosis and microRNAs.**

Of the many CVDs' affecting human health today, atherosclerosis, a systemic disorder characterized by the narrowing of arteries, serves as one of the largest risk factors for mortality as it is the underlying cause for majority of clinical cardiovascular events including myocardial infarction and stroke<sup>2</sup>. The accumulation of oxidized phospholipids within the vessel wall due to oxidation of LDL has long been considered one of the key primary events in the development of the atherosclerotic plaque. In endothelial cells (ECs), exposure to oxidized phospholipids promotes the transition from a quiescent state to a pro-inflammatory phenotype characterized by increased vascular permeability and expression of adhesion molecules<sup>3</sup>. These changes lead to enhanced monocyte recruitment into the vessel wall, thus promoting lesion development and are characterized by dramatic alteration in gene expression.

MiRNAs are a highly conserved species of non-coding RNAs which act to regulate gene expression through target mRNA degradation or impaired translation. MiRNAs are hypothesized to be a common feature of gene expression regulation throughout the animal kingdom with the identification of novel miRNAs having been reported in viruses, plants, single cell organisms and large mammals such as humans. In humans, dysregulation of miRNA expression has been linked to the development of numerous diseases including cancer, neurological disorders and

CVDs<sup>4,6</sup>. In chapter 2's "Regulation of NF- $\kappa$ B signaling by oxidized glycerophospholipid and IL-1 $\beta$  induced miRs-21-3p and -27a-5p in human aortic endothelial cells" I identified and characterized the function of the novel miRNAs miR-21-3p and -27a-5p in ECs. I have shown that these two miRNAs are upregulated in response to inflammatory stimuli such as oxidized lipids and cytokines and act to modulate the extent of NF- $\kappa$ B activation in ECs. In a broader context, this project identified two novel miRNA regulators of NF- $\kappa$ B signaling in endothelial cells. The NF- $\kappa$ B family of transcription factors are vital to the regulation of numerous cellular processes including the immune and stress response, cell adhesion, and protection against apoptosis<sup>7</sup>. In the cardiovascular system, dysregulation of NF- $\kappa$ B signaling is crucial in the development of both atherosclerosis and heart failure<sup>8,9</sup>

This project, in particular is set apart from the others included in my dissertation as it on its own it is not a true systems genetic project. It does not utilize genetic variation a tool for discovery of novel biological features relevant to the disease model even though it does utilize multiple global profiling technologies (RNA sequencing and microarrays) in attempt to dissect the fundamental research questions. Rather, this project exemplifies how systems genetics discoveries can further drive science by asking critical questions about previous findings. This study was driven by attempt to explain the gene co-expression network generated from an earlier systems genetics analysis of the response of 147 endothelial cell donors to oxidized phospholipids<sup>10</sup>. If there is one project that has to define my graduate training, it was this project that had the most significant impact on me during my training. I take an immense pride in the amount of work I contributed to this project; from the initial experimental design, to learning the multiple molecular biology and bioinformatics techniques utilized during the course of the study to the preparation of manuscript. I grew as both a person and scientist during the 3 years I spent

generating the data and I look back on this period as vital in shaping my ability to tackle the projects I worked on during the later portions of my graduate career.

### **The role of *Abcc6* in cardiac fibrosis and calcification.**

Cardiac fibrosis is defined as the accumulation of fibrotic tissue within the myocardium. Uncontrolled fibrogenesis due to chronic stress or cardiac damage has a significant impact on the biophysical properties of the myocardium leading to impaired cardiac contractility and conductivity. Consequently, the development of cardiac fibrosis is fundamental to progression of numerous cardiac disorders including congestive heart failure (CHF). However, unlike other risk factors for CHF, such as hypertension and diabetes, our understanding of the role of genetic variation as a risk factor for developing cardiac fibrosis remains limited. In chapter 3 “Mapping genetic contributions to cardiac pathology induced by Beta-adrenergic stimulation in mice.” I present a project involving the identification and validation of the ABC transporter, *Abcc6* as novel contributor to the development of Isoproterenol (ISO) induced cardiac fibrosis. As part of this project I applied system genetics analysis techniques to identify key candidates residing within the six isoproterenol fibrosis loci in addition to assisting in the *Abcc6* validation experiments and manuscript preparation.

In mouse, *Abcc6* was first characterized in the Lusis lab over 10 years ago as a gene contributing to the development of dystrophic cardiac calcification<sup>11</sup>. In humans, loss of function mutations in *Abcc6* cause the connective tissue disorder Pseudoxantoma Elasticum, which is characterized premature blindness and early atherosclerosis due to soft tissue calcification<sup>12-14</sup>. Nonetheless, very little is known on the exact biological processes altered by loss of *Abcc6* that contribute to the abnormal soft tissue calcification seen in both mice and humans lacking functional ABCC6 protein.

The accumulation of extracellular pyrophosphate (PPi) is a potent inhibitor of calcification of soft tissues. Extracellular PPi is generated through the NPP1-mediated hydrolysis of extracellular NTPs and/or by the intracellular-to-extracellular channeling of PPi by the transmembrane protein ANK<sup>15</sup>. In recent work implicating a role for *Abcc6* in mediating the release of nucleotide triphosphates (NTP) into the extracellular matrix, it has been shown that loss of *Abcc6* in mice causes > 40% decrease in circulating pyrophosphate levels in the plasma. These findings strongly suggest that decreased levels of circulating pyrophosphate as a possible mechanism for soft tissue calcification seen in PXE, yet many questions remain about how loss of *Abcc6* leads to the observed decreased levels circulating pyrophosphate.

However, recent work done in collaboration with Arjun Deb here at UCLA during my PhD hopes to provide evidence towards a molecular mechanism of how loss of *Abcc6* leads to soft tissue calcification. The Deb lab has recently shown that cardiac fibroblasts upon injury adopt an osteogenic like fate that contributes to the development of tissue calcification. Furthermore, they have found that this contribution towards calcification is cell autonomous as fibroblasts from an injured heart implanted in uninjured mice will still form calcific modules. During the course of these studies they also found that fibroblasts from the inbred strain C3H/HeJ had greater calcification potential than those C57BL/6J, which in part is due to loss of *Abcc6* in C3H/HeJ strain.

Therefore as part of this collaboration I have investigated the relationship between *Abcc6* and 39 markers of osteogenesis using the HMDP. Using this approach we identified a strong negative correlation between *Abcc6* and Alkaline Phosphatase (*Alpl*) the enzyme that converts extracellular PPi to phosphate. Using adenoviral over-expression of *Abcc6* in fibroblasts, a cell type crucial to wound healing, the Deb lab has shown that there is decreased *Alpl* activity upon

Abcc6 overexpression. Furthermore, we have found a strong positive correlation between expression of the gene *Abcc6* and *Enpp1*, one of the enzymes responsible for the production of extracellular PPI, suggesting that loss of Abcc6 may lead to decreased *Enpp1* as possible mechanism for the observed decrease in extracellular PPI levels. While this collaboration is still in its early stages I strongly believe we have made significant findings towards identifying the molecular mechanisms behind *Abcc6* deficiency and soft tissue calcification with these studies.

### **The role of *Myh14* in regulation of left ventricular mass as determined by serial echocardiography.**

In the last four decades, the use of small-animal models such as mice to study complex cardiovascular pathophysiology has proven to be invaluable<sup>16</sup>. Nonetheless, important considerations must be taken when translating findings in mouse models to humans due to the vast differences in their cardiovascular systems and clinical limitations for phenotyping<sup>17</sup>. For example, cardiac remodeling is one of the key prognostic determinants of clinical HF in human patients. However, in humans, unlike animal models, clinicians are limited to the use of echocardiography to determine the extent of cardiac remodeling following cardiac injury.

In chapter 4 “Genetic Dissection of Cardiac Remodeling in an Isoproterenol-induced Heart Failure Mouse Model” I present a project involving the validation and characterization of the gene *Myh14* as involved with cardiac remodeling as determined by serial echocardiography in mouse. *Myh14* is a non-muscle myosin involved in mechanotransduction with no obvious phenotype at baseline (Jessica Reference 2). As part of this project I characterized the specific role of *Myh14* in cardiomyocytes *in vitro* using siRNA knockdown of *Myh14* in neonatal rat ventricular myocytes (NRVM). I showed that decreased expression of *Myh14* causes NRVMs to display poor attachment to tissue culture plates. Furthermore, cells with decreased expression of

*Myh14* show impaired ability to undergo cellular hypertrophy and display decreased cell viability upon stimulation with phenylephrine or ISO.

To elucidate possible molecular mechanisms for the observed effects of *Myh14* deficiency in left ventricular mass hypertrophy, I identified those genes that were strongly correlated to cardiac *Myh14* expression and performed an eQTL co-mapping analysis. In this approach, we first identify SNPs which reside nearby *Myh14* and are predicted to regulate *Myh14* expression, so called cis-eQTLs. We then identify those transcripts that are significantly correlated to *Myh14* whose expression map to previously identified SNPs predicted to regulate *Myh14* expression. We therefore can then hypothesize that the transcripts whose expression maps to the *Myh14* cis-eQTL are regulated by *Myh14*. Using this approach I found that the cis-eQTL locus that controls *Myh14* expression in heart also controls the expression of transcription factor, *Foxo1* (p-value =  $1.20 \times 10^{-5}$ ) and cell cycle regulator *Cdk11b* (p-value =  $9.08 \times 10^{-6}$ ). These two genes serve as initial targets for further investigation into the exact molecular mechanism behind *Myh14*'s effect on left ventricular mass.

### ***Adamts2* a novel regulator of cardiac hypertrophy**

Congestive heart failure is a complex disease characterized by impaired cardiac function leading to the development of peripheral edema, shortness of breath, and fatigue. Despite the success of genome wide association studies (GWAS) in other CVDs such as atherosclerosis, atrial fibrillation and hypertension, attempts to dissect the genetic contribution towards the development of CHF have had limited success. Network modeling approaches, in contrast to GWAS emphasize analysis of multiple interactions across a dataset to identify novel interactions related to the trait of interest. In chapter 5 “A systems genetics approach to identify genetic pathways and key drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in

mice.” I present the identification and characterization of the gene *Adamts2* as a novel regulator of cardiac hypertrophy in vitro. Utilizing a combination of gene co-expression network modeling and molecular biology techniques I identified 5 new targets of *Adamts2*, *Kcnv2*, *Mfap2*, *Tnc*, *Nppa* and *Nppb*. In addition I showed that *Adamts2* is a novel regulator of isoproterenol induced cardiac hypertrophy in vitro.

As part of a collaboration with Ionis Pharmaceuticals, I am in the preliminary phases of testing a novel antisense oligo to *Adamts2* in vivo. However, when translating the earlier in vitro *Adamts2* findings into an in vivo application there are a few key considerations. First, the in vitro validation was performed using a cell culture model that emphasizes a shorter-term isoproterenol treatment when compared to the standard 21 day isoproterenol treatment used for the HMDP. Furthermore, the cell culture model does not integrate the crosstalk between cardiomyocytes and other isoproterenol-responsive tissues, such as adipose. Finally, as loss of *Adamts2* is known to contribute to rare connective tissue disorder Ehlers Danlos Syndrome, whose clinical manifestations include joint hypermobility and aortic dissection, care must be taken in managing possible secondary effects of *Adamts2* inhibition.

Therefore, as part of these preliminary trials, I have begun injecting mice weekly with the *Adamts2* ASO at spectrum of doses followed by treatment the mice with isoproterenol for 21 days. I will determine efficacy of the dose by comparing phenotypic changes such as total heart weight, liver and lung weight, with animals treated with saline as a negative control for the ASO in addition to performing qPCR to determine percent knockdown of *Adamts2*. In parallel I will also measure the toxicity effects of *Adamts2* ASO treatment in combination with Isoproterenol treatment by expression of liver ALT/AST enzymes through ELISA. My goal is to find an

effective dose of the ASO that limits the toxicity but shows strong effects on isoproterenol-induced hypertrophy.

Congestive Heart Failure (CHF) remains a growing global public health epidemic, with an estimated 5.7 million patients in the US and global prevalence of 37.7 million<sup>18</sup>. Furthermore, epidemiological predictions strongly imply that hospitalizations for CHF expected to rise due to an increasing ageing population, improved survival after other types of cardiovascular disease, and reductions in rates of sudden cardiac death<sup>18</sup>. As such improved care and prevention for CHF is a growing necessity, however many large phase III trials of HF drugs have consistently fallen short of expectations, either failing to show clear efficacy or raising safety concerns<sup>19</sup>. A number of factors have been implicated in this failure to develop novel CHF therapeutics, such as studying the 'right' patient populations, targeting the relevant pathways that have a potential to improve outcomes and using appropriate surrogate outcomes that are correlated with hard end points. Nonetheless, I strongly believe that the systems genetics analysis possible with HMPD will help identify future novel drug targets such as *Adamts2* for improvement in CHF patient care.

### **From mouse to man – Systems Genetic Approaches in Humans and Precision Medicine**

Despite the success of reductionist, single approaches in identifying molecular mechanisms leading to the development cardiovascular disorders, current estimates predict that 30 million deaths annually will be due to CVDs across the globe by the year 2030. CVDs such as atherosclerosis and CHF are complex multi-systemic disorders, characterized by numerous environmental and genetic risk factors whose interactions within a given individual predispose that person to a greater risk of disease than the individual sum of the risk factors in consideration. As such, understanding the interactions of environmental risk factors within across multiple

unique genetic backgrounds is crucial to improving our understanding CVDs. Systems genetics, a holistic analysis framework for understanding the flow of biological information through integration of multi-layered –omics profiling, mathematical modeling and molecular biology, is highly suited to such complex biological disorders as atherosclerosis and CHF. Nonetheless, while animal models, such as the HMDP serve as a strong starting point for systems genetics mediate discovery, the ultimate goal of disease driven biomedical research is translation into patients.

On January 20<sup>th</sup>, 2015, during his State of the Union address, President Obama announced the Precision Medicine Initiative, a bold, new patient-powered research initiative to revolutionize how we treat disease and improve patient care<sup>20</sup>. Precision medicine is the practice of using a combination of binary (sex, genotype) and continuous (age, bodyweight) factors to bin patients into precise categories to facilitate evidence-based decision regarding a given patients treatment plan<sup>21</sup>. With its decreasing cost and improving data analysis pipelines genetic information is becoming a powerful tool in the hands of not only researchers, but clinicians and patients. The FDA now includes pharmacogenetics information in their labeling for numerous drugs to improve patient treatment. Clinicians are now regularly utilizing exome sequencing to identify disease causing genes in rare mendelian disorders who were previously of unknown genetic etiology. Furthermore, patients themselves are now in control of accessing own genetic information through use of such whole genome genotyping services such 23andMe. It is becoming increasingly obvious that the 21<sup>st</sup> century is an era where systems genetics analysis in large human cohorts will now be realistically feasible due to the ease of access to genetic information. While this is an exciting time, our ability to translate an individual's genetic profile into actionable risk management and patient care for common diseases is still many years away.

## References

- 1 Wong, N. D. Epidemiological studies of CHD and the evolution of preventive cardiology. *Nat Rev Cardiol* **11**, 276-289, doi:10.1038/nrcardio.2014.26 (2014).
- 2 Lusis, A. J. Atherosclerosis. *Nature* **407**, 233-241, doi:10.1038/35025203 (2000).
- 3 Lee, S. *et al.* Role of Phospholipid Oxidation Products in Atherosclerosis. *Circulation Research* **111**, 778-799, doi:10.1161/circresaha.111.256859 (2012).
- 4 Di Leva, G., Garofalo, M. & Croce, C. M. MicroRNAs in Cancer. *Annual Review of Pathology: Mechanisms of Disease* **9**, 287-314, doi:10.1146/annurev-pathol-012513-104715 (2014).
- 5 Wang, W., Kwon, E. J. & Tsai, L. H. MicroRNAs in learning, memory, and neurological diseases. *Learning & Memory* **19**, 359-368, doi:10.1101/lm.026492.112 (2012).
- 6 Quiat, D. & Olson, E. N. MicroRNAs in cardiovascular disease: from pathogenesis to prevention and treatment. *Journal of Clinical Investigation* **123**, 11-18, doi:10.1172/jci62876 (2013).
- 7 Smahi, A. *et al.* The NF-kappaB signalling pathway in human diseases: from incontinentia pigmenti to ectodermal dysplasias and immune-deficiency syndromes.
- 8 de Winther, M. P. J. Nuclear Factor B Signaling in Atherogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **25**, 904-914, doi:10.1161/01.atv.0000160340.72641.87 (2005).
- 9 Gordon, J. W., Shaw, J. A. & Kirshenbaum, L. A. Multiple Facets of NF-κB in the Heart: To Be or Not to NF-κB. *Circulation Research* **108**, 1122-1132, doi:10.1161/circresaha.110.226928 (2011).

- 10 Romanoski, C. E. *et al.* Network for activation of human endothelial cells by oxidized phospholipids: a critical role of heme oxygenase 1. *Circ Res* **109**, e27-41, doi:10.1161/CIRCRESAHA.111.241869 (2011).
- 11 Meng, H. *et al.* Identification of Abcc6 as the major causal gene for dystrophic cardiac calcification in mice through integrative genomics. *Proceedings of the National Academy of Sciences* **104**, 4530-4535, doi:10.1073/pnas.0607620104 (2007).
- 12 Bergen, A. A. *et al.* Mutations in ABCC6 cause pseudoxanthoma elasticum.
- 13 Le Saux, O. *et al.* Mutations in a gene encoding an ABC transporter cause pseudoxanthoma elasticum.
- 14 Ringpfeil, F., Lebowitz M. G. - Christiano, A. M., Christiano Am Fau - Uitto, J. & Uitto, J. Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a transmembrane ATP-binding cassette (ABC) transporter. doi:D - NLM: PMC18548 EDAT- 2000/05/17 09:00 MHDA- 2000/07/08 11:00 CRDT- 2000/05/17 09:00 AID - 10.1073/pnas.100041297 [doi] AID - 100041297 [pii] PST - ppublish.
- 15 Stefan, C., Jansen, S. & Bollen, M. NPP-type ectophosphodiesterases: unity in diversity. *Trends Biochem Sci* **30**, 542-550, doi:10.1016/j.tibs.2005.08.005 (2005).
- 16 Patten, R. D. & Hall-Porter, M. R. Small Animal Models of Heart Failure: Development of Novel Therapies, Past and Present. *Circulation: Heart Failure* **2**, 138-144, doi:10.1161/circheartfailure.108.839761 (2009).
- 17 Zaragoza, C. *et al.* Animal Models of Cardiovascular Diseases. *Journal of Biomedicine and Biotechnology* **2011**, 1-13, doi:10.1155/2011/497841 (2011).
- 18 Ziaeeian, B. & Fonarow, G. C. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol*, doi:10.1038/nrcardio.2016.25 (2016).

- 19 Vaduganathan, M., Greene, S. J., Ambrosy, A. P., Gheorghiade, M. & Butler, J. The disconnect between phase II and phase III trials of drugs for heart failure. *Nature Reviews Cardiology* **10**, 85-97, doi:10.1038/nrcardio.2012.181 (2013).
- 20 The White House, O. o. t. P. S. *FACT SHEET: President Obama's Precision Medicine Initiative*, <<https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative>> (2015).
- 21 Insel, P. A., Amara, S. G. & Blaschke, T. F. Introduction to the theme "Precision medicine and prediction in pharmacology". *Annu Rev Pharmacol Toxicol* **55**, 11-14, doi:10.1146/annurev-pharmtox-101714-123102 (2015).